Va 2

NCS TIB 91-3

NATIONAL COMMUNICATIONS SYSTEM

TECHNICAL INFORMATION BULLETIN 91-3

DEVELOPMENT OF A FEDERAL STANDARD FOR A VIDEO CODEC OPERATING AT P X 64 Kbps

FEBRUARY 1991

OFFICE OF THE MANAGER
NATIONAL COMMUNICATIONS SYSTEM

WASHINGTON, D.C. 20305

REPORT DOCUMENTATION PAGE

1 AGENCY USE ONLY

February 1991 | Final

4 TITLE AND SUBTITLE

Development of a Federal Standard for a Video Codec
Operating at P X 64 Kbps

C-DCA100-87-C-0078

6. AUTHOR(S)

7 PERFORMING

Delta Information Systems, Inc.
300 Welsh Road, Suite 120
Horsham, PA 19044-2273

9 SPONSORING

National Communictions System
Office of Technology & Standards
701 S. Court House Road
Arlington, VA 22204-2199

NCS TIB 91-3

Approved for Public Release; distribution is unlimited.

This document is a final report under the contract entitled "Development of Federal Telecommunication Standards relating to Digital Facsimile and Video Teleconferencing" and "Development of a Federal Standard for a Vieo Codec operating at P X 64 Kbps." There has been a major effort to develop a standard for a video teleconferencing/ video telephone terminal. One major thrust is that the audio visual terminal be compatible with the ISDN. Since the basic building block of the ISDN is the B channel operating at 64 Kbps, the generic term "P X 64 Kbps" refers to operation of this audio visual terminal at integral values of P up to a maximum of 30.

Video teleconferencing | ISDN | 218
Video telephone
Digital Facsimile

Unclassified | Unclassified | Unclassified | Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines to meet optical scanning requirements.**

Block 1. Agency Use Only (Leave Blank)

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Names(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency. Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of ..., To be published in .... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availablity Statement. Denote public availability or limitation. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR)

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - DOD - Leave blank
- **DOE** - DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports
- **NASA** - NASA - Leave blank
- **NTIS** - NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

NCS TECHNICAL INFORMATION BULLETIN 91-3

DEVELOPMENT OF A FEDERAL STANDARD FOR A VIDEO CODEC OPERATING AT PX64 KBPS

JANUARY 1991

PROJECT OFFICER

GARY M. REKSTAD
Electronics Engineer
Office of NCS Technology
and Standards

APPROVED FOR PUBLICATION:

Dennis Bodson

DENNIS BODSON
Assistant Manager
Office of NCS Technology
and Standards

FOREWORD

Among the responsibilities assigned to the Office of the Manager, National Communications System, is the management of the Federal Telecommunication Standards Program. Under this program, the NCS, with the assistance of the Federal Telecommunication Standards Committee identified, develops, and coordinates proposed Federal Standards which either contribute to the interoperability of functionally similar Federal telecommunication systems or to the achievement of a compatible and efficient interface between computer and telecommunication systems. In developing and coordinating these standards, a considerable amount of effort is expended in initiating and pursuing joint standards development efforts with appropriate technical committees of the International Organization for Standardization, and the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union. This Technical Information Bulletin presents and overview of an effort which is contributing to the development of compatible Federal, national, and international standards in the area of teleconferencing. It has been prepared to inform interested Federal activities of the progress of these efforts. Any comments, inputs or statements of requirements which could assist in the advancement of this work are welcome and should be addressed to:

Office of the Manager
National Communications System
ATTN: NCS-TS
Washington, DC 20305-2010

DEVELOPMENT OF A FEDERAL STANDARD FOR A VIDEO CODEC OPERATING AT P x 64 KBPS

February, 1991

FINAL REPORT
DCA100-87-C-0078
TASK ORDER NUMBER 89-5

Submitted to:
NATIONAL COMMUNICATIONS SYSTEM
WASHINGTON, DC

DELTA INFORMATION SYSTEMS, INC.
300 Welsh Road, Ste. 120
Horsham, PA 19044-2273
TEL: (215) 657-5270 FAX: (215) 657-5273

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1.0 INTRODUCTION

This document summarizes work performed by Delta Information Systems, Inc. (Delta) for the National Communications System (NCS), Office of Technology and Standards. This office is responsible for the management of the Federal Telecommunications Standards Program, which develops telecommunications standards, whose use is mandatory for all Federal departments and agencies.

This document is a final report for Task Order 89-5 on Contract DCA100-87-C-0078. The titles for the contract and Task Order are listed below.

- Contract DCA100-87-C-0078
  Development of Federal Telecommunication Standards Relating to Digital Facsimile and Video Teleconferencing
- Task Order 89-5
  Development of a Federal Standard for a Video Codec operating at P x 64 Kbps

For several years, there has been a major effort by the CCITT and by ANSI (American National Standard Institute) to develop a standard for a video teleconferencing/video telephone terminal. One major thrust of this activity is that the audio visual terminal be compatible with the ISDN. Since the basic building block of the ISDN is the B channel operating at 64 Kbps, the generic term "P x 64 Kbps" refers to operation of this audio visual terminal at integral values of P up to a maximum of 30. Values of P which are of greatest interest are 1, 2, 6, 12, 24, and 30.

Work on the P x 64 standard during this report interval (June 15, 1989 to December 15, 1990) was extremely fruitful. Highlights are listed below.

- There were two meetings of the Specialist's Group on Visual Telephony Coding resulting in a draft standard of Recommendation H.261 entitled "Video Codec for Audiovisual Services at P x 64 Kbit/s.
- CCITT Study Group XV, Working Group I met in July, 1990 to finalize the drafts for the five P x 64 Recommendations listed below.
  H.261 Video codec for audiovisual services at P x 64 kbit/s

H.221	Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices
H.242	System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s
H.230	Frame synchronous control and indication signals for audiovisual systems
H.320	Narrow-band visual telephone systems and terminal equipment

- These five P x 64 standards became official CCITT Recommendations on December 14, 1990. A copy of the final Recommendations is included in Appendix A.
- The NCS has developed a draft of a Federal Standard (Number 1080) which is based upon Recommendation H.261. A copy of this draft is included in Appendix B.

Section 2.0 of this report summarizes the P x 64 standardization activity by Delta during the report interval. Section 3.0 provides an overview of all P x 64 Recommendations other than H.261. Since Recommendation H.261 is particularly critical to the P x 64 standard, it is discussed in more detail in Section 4.0. Conclusions are drawn in Section 5.0.

2.0 STANDARDIZATION ACTIVITY

Figure 2.1 illustrates the relationship between those organizations which actively participate in the P x 64 standardization process. The CCITT is a part of the United Nations, and its purpose is to develop formal "Recommendations" to insure worldwide communications are accomplished efficiently and effectively. The CCITT works in four-year cycles, and at the end of each period a complete set of Recommendations is published. The "Red Books" and "Blue Books" containing these Recommendations were published in 1984 and 1988 respectively.

In the 1984 Red Books the first Recommendations for a teleconferencing codec (H.120 and H.130) were established. These Recommendations were defined specifically for the European region (625 lines; 2.048 Mbit/s primary rate) and for interconnection between Europe and other regions. Since no Recommendation existed for non-European regions it lacked true international scope, and in 1984 the CCITT established a "Specialists Group on Coding for Visual Telephony" to develop a truly international Recommendation. Figure 2.1 illustrates the location of this group in the CCITT hierarchy. The CCITT established two objectives for the Specialists Group: (1) to develop a Recommendation for a video codec for teleconferencing application operating at the bit rates of Nx384 Kbit/s (N=1 through 5), and (2) to begin the standardization process for a video codec for teleconferencing/ video telephone application operating at bit rates of Mx64 Kbit/s (M=1,2).

The Specialists Group met seventeen times from 1984 through 1989. Representatives from Canada, Finland, France, Germany, Italy, Japan, Korea, Netherlands, Norway, Sweden, the U.K., and the United States were members of the committee.

At the September meeting in 1988, it was determined that the compression algorithm chosen for Nx384 Kbit/s was sufficiently flexible that it could be extended, with good performance, down to 64 Kbit/s. At that time, the Specialists Group shifted their focus to develop a single Recommendation to code video at all bit rates from 64 Kbit/s to 2 Mbit/s; i.e. to code at rates of px64 Kbit/s, where the key values of p are 1, 2, 6, 24, and 30.

In 1989, a number of organizations in Europe, U.S. and Japan developed flexible codec systems to meet a preliminary specification of the standard. Various systems were interconnected in the laboratory and by long distance communication channels to validate the Recommendation. These tests were highly

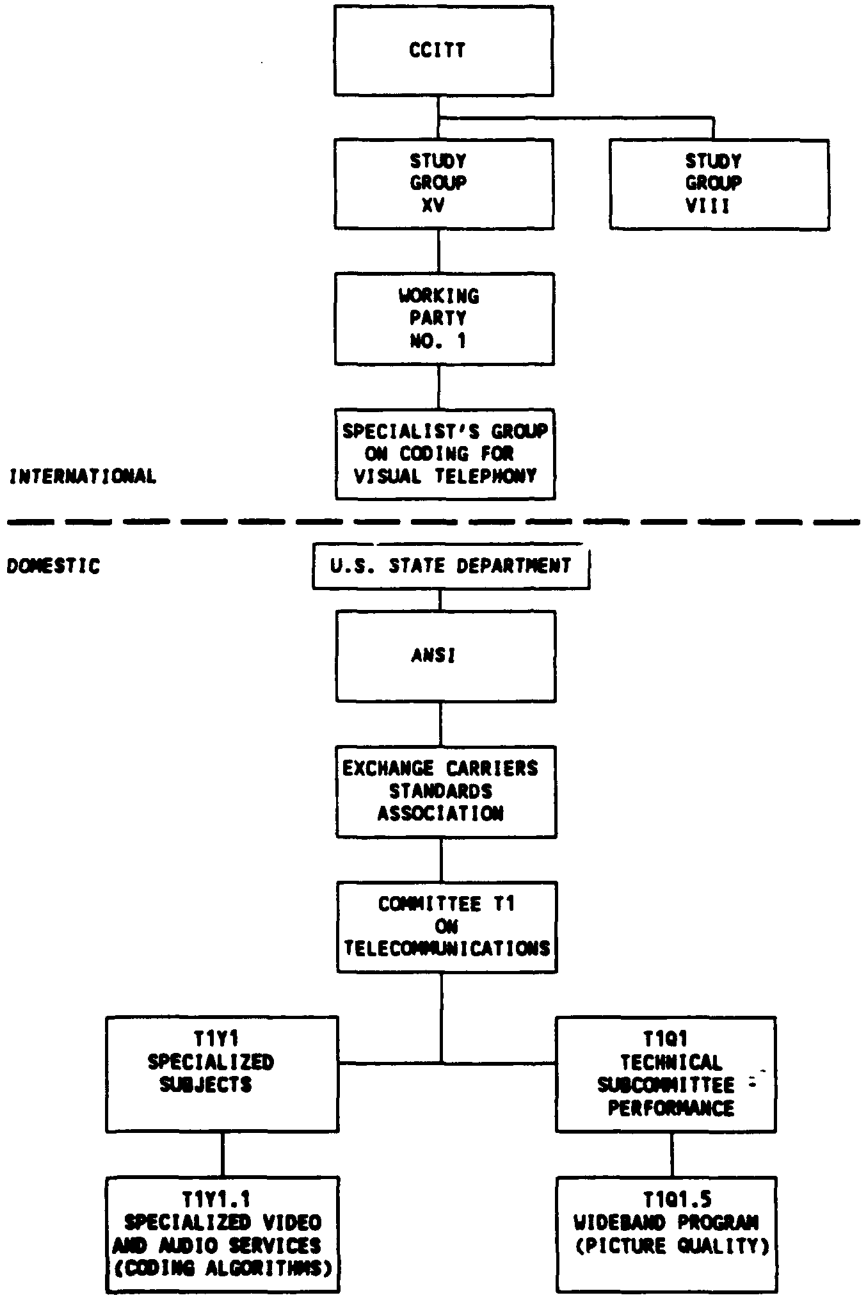


FIGURE 2.1 ORGANIZATIONS CONTRIBUTING TO STANDARDS FOR VIDEO TELECONFERENCING

successful and encouraging. A preliminary version of the final H.261 Recommendation appears in the recent CCITT Blue Book. However, this is incomplete, and a final draft of a more complete H.261 was written and submitted to the CCITT for approval by means of a new Accelerated Procedure. The revised H.261 Recommendation was formally approved by the CCITT on December 14, 1990.

Figure 2.1 also illustrates those domestic standards organizations which develop the U.S. technical positions on issues related to video teleconferencing and video telephone. The T1 committees, which are accredited by ANSI (American National Standards Institute), work on two different aspects of teleconferencing: the T1Y1.1 committee is responsible for the coding algorithms, while T1Q1.5 is responsible for defining and measuring the quality of service to be provided by teleconference systems.

Delta personnel attended a wide range of meetings of standards organizations during the report interval (June 15, 1989 to December 15, 1990). The meetings attended are listed below.

T1Y1.1

- July 12-89 — Denver, Colorado
- April 3,4-90 — Dallas, Texas
- July 11-90 — Seattle, Washington
- October 3-90 — Rockville, Maryland
- October 28,29-90 — Chicago, Illinois
  Joint Meeting of T1Y1.1, T1Q1.5, T1S1

CCITT

- June 12-15, 1989 — Stuttgart, Germany
  Specialists Group
  Review of Field Tests
- October 23-26, 1989 — Ipswich, England
  Editorial Subgroup of Specialists Group

- November 7-10, 1989 Tokyo, Japan
  Specialists Group
  Finalized Draft H.261 Recommendation

- July 16-20, 1990 Geneva, Switzerland
  Study Group XV, Working Party I
  Finalized Draft P x 64 Recommendations

STATE DEPARTMENT STUDY GROUP C

- June 19, 1990 Washington, DC
  Developed the U.S. Contribution for July, 1990 SG XV/I Meeting

MIL STD 188B

The Department of Defense is developing Standard MIL STD 188B for teleconferencing. Delta personnel met with DoD representatives to contribute to this standard.

- September 26, 1989 Defense Communication Engineering Center
  Reston, Virginia

- September 27, 1990 Ft. Monmouth, New Jersey

3.0 OVERVIEW OF P x 64 RECOMMENDATIONS

As explained in the above sections P x 64 is a generic term which encompasses a broad class of audiovisual services defined in CCITT Recommendations H.320, H.261, H.221, H.242, and H.230. Basically, any audio visual terminal equipment must rigorously adhere to all of these Recommendations (including the I.400 series which defines ISDN operations) to be interoperable. Since the H.320 Recommendation describes the interrelationship between all of the P x 64 Recommendations it will be described first. The H.261 is the cornerstone of the P x 64 Recommendations, and is described in a Section 4.0.

3.1 H.320 Narrow-band Visual Telephone Systems and Terminal Equipment

Recommendation H.320 contains the diagram in Figure 3.1 which illustrates the relationship between all of the P x 64 Recommendations. The only element in this diagram which may not be obvious is the input from "Telematic Equipment". In this case "telematic equipment" refers primarily to sources of still pictures. The H.221 Recommendation specifically makes provision for multiplexing still picture signals with video and audio signals. The still picture signals which can be accepted also must conform to standards which either exist or are in development. There are three general types of still picture signals which can be handled by a P x 64 terminal.

1. <u>ISO (International Standards Organization) Still Picture Standard</u>

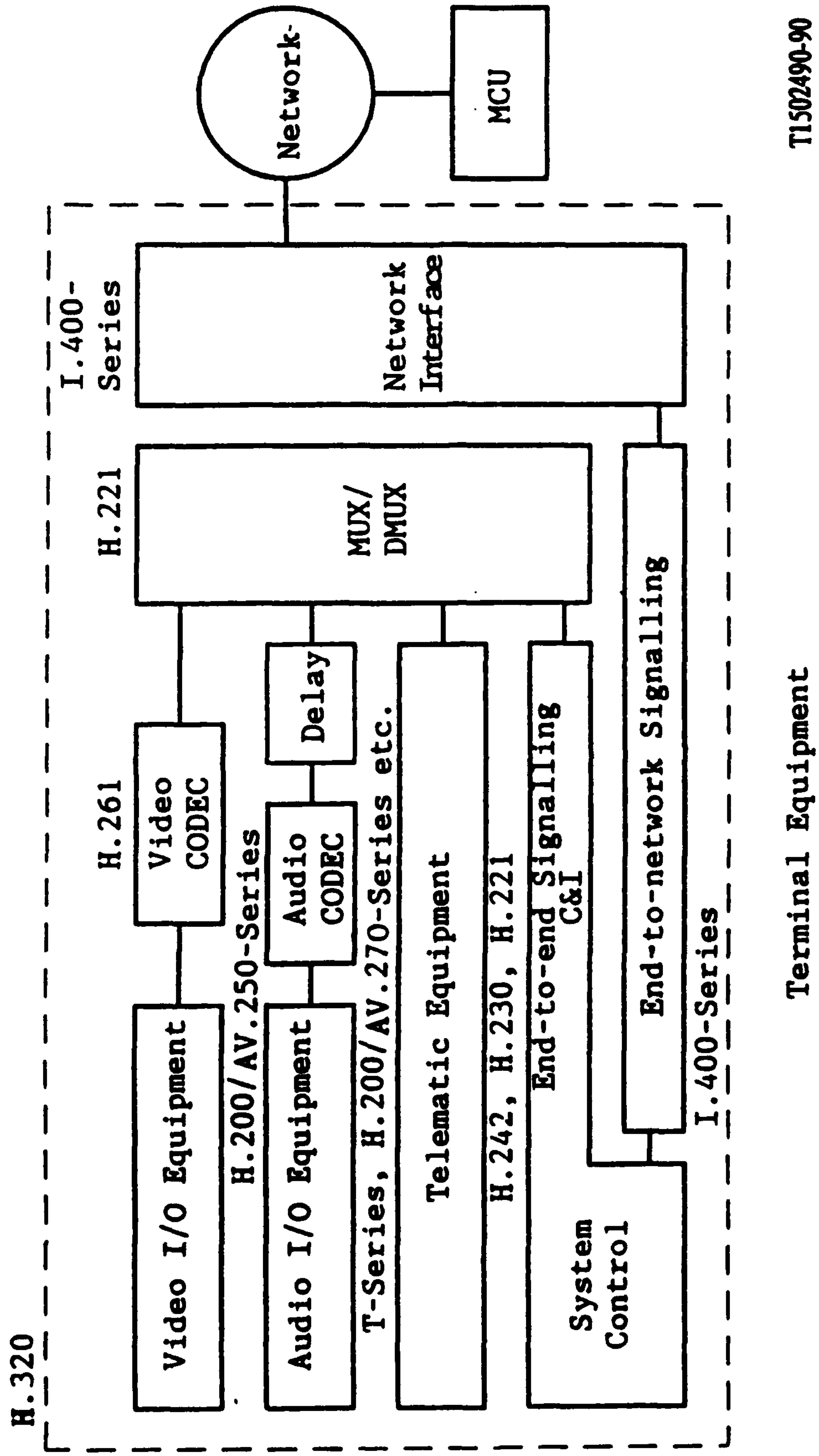
The ISO and CCITT have established a Joint Photographic Experts Group (JPEG) to develop a general purpose international standard for the coding of continuous tone (gray scale or color) images. Although this standard is in the draft status, it is well defined and will prove to be a valuable way of communicating graphics in P x 64/H.320 terminals.

2. <u>Group 3 Facsimile</u>

Group 3 facsimile is one of the most successful standards ever developed. It has keyed the fax revolution. The H.320 terminal can handle signals conforming to the CCITT Group 3 standards (T.4, T.30).

3. <u>Group 4 Facsimile</u>

Several Group 4 facsimile standards have been finalized (T.563, T.5, T.503,



Terminal Equipment

MCU: Multipoint Control Unit

FIGURE 1/H.320

__Visual telephone system__

FIGURE 3.1

T.6) and some are still in development. The H.320 audiovisual terminal can handle Group 4 facsimile signals.

One function of the H.320 Recommendation is to define the phases of establishing a visual telephone call as listed below.

Phase A:	Call set-up, out-band signalling
Phase B1:	Mode initialization on initial channel
Phase CA:	Call set-up of additional channel(s), if relevant
Phase CB1:	Initialization on additional channel(s)
Phase B2 (or CB2):	Establishment of common parameters
Phase C:	Visual telephone communication
Phase D:	Termination phase
Phase E:	Call release

Another function of Recommendation H.320 is the definition of 16 different types of visual telephone terminals and their modes of operation (see Figures 3.2 and 3.3 respectively).

3.2 H.221 Frame Structure for a 64 to 1920 Kbit/s Channel in Audiovisual Teleservices

The purpose of this Recommendation is to define a frame structure for audiovisual teleservices in single or multiple B or HO channels or a single H11 or H12 channel which makes the best use of the characteristics and properties of the audio and video encoding algorithms, of the transmission frame structure, and of the existing CCITT Recommendations. It offers several advantages:

- It is simple, economic and flexible. It may be implemented on a simple microprocessor using well-known hardware principles.
- It is a synchronous procedure. The exact time of a configuration change is the same in the transmitter and the receiver. Configurations can be changed at 20 ms intervals.
- It needs no return link for audiovisual signal transmission, since a configuration is signalled by repeatedly transmitted codewords.
- It is very secure in case of transmission errors, since the code controlling the

TABLE 2/H.320

Visual telephone terminal type

Mode	Type X (Note 2)										Type Y (Note 3)					Type Z	
	a	b1	b2	b3	b4	b5	c	d	e	f	1	2	3	4	5	α	β
a0 B(audio only)	X	X	X	X	X	X	X	X	X	X							
a1 B(H.200/AV.254 audio)	X	X	X	X			X	X	X	X							
b1 2B(G.711 audio)		X	X	X	X	X	X	X	X	X							
b2 2B(G.722 audio)			X	X		X	X	X	X	X							
b3 2B(H.200/AV.254 audio)				X			X	X	X	X							
c 3B							X	X	X	X							
d 4B								X	X	X							
e 5B									X	X							
f 6B										X							
g H0											X	X	X	X	X		
h 2H0												X	X	X	X		
i 3H0													X	X	X		
j 4H0														X	X		
k H11																X	
l 5H0															X		
m H12																	X

Note 1 - "X" means the mode is equipped with the terminal of the type.

Note 2 - Types Xb4 and Xb5 are defined to take into account that H.200/AV.254 has not yet been recommended.

Note 3 - Terminal of this type must have the H0-6B compatible mode defined in Recommendation H.221.

FIGURE 3.2

TABLE 1/H.320

<u>Communication modes of visual telephone</u>

<table>
<tr><th colspan="2" rowspan="2">Visual Telephone Mode (Note 1)</th><th rowspan="2">Channel Rate (kbit/s)</th><th rowspan="2">ISDN Channel (Note 2)</th><th colspan="2">ISDN Interface</th><th colspan="2">Coding</th></tr>
<tr><th>Basic</th><th>Primary Rate</th><th>Audio</th><th>Video</th></tr>
<tr><td rowspan="2">a</td><td>a0</td><td rowspan="2">64</td><td rowspan="2">B</td><td colspan="2" rowspan="5"></td><td>G.711</td><td>not applicable</td></tr>
<tr><td>a1</td><td>H.200/AV.254</td><td rowspan="15">H.261</td></tr>
<tr><td rowspan="3">b</td><td>b1</td><td rowspan="3">128</td><td rowspan="3">2B</td><td>G.711</td></tr>
<tr><td>b2</td><td>G.722</td></tr>
<tr><td>b3</td><td>H.200/AV.254 /253 (Note 1)</td></tr>
<tr><td colspan="2">c</td><td>198</td><td>3B</td><td rowspan="11">not applicable</td><td rowspan="11">applicable</td><td rowspan="11">G.722</td></tr>
<tr><td colspan="2">d</td><td>256</td><td>4B</td></tr>
<tr><td colspan="2">e</td><td>320</td><td>5B</td></tr>
<tr><td colspan="2">f</td><td>384</td><td>6B</td></tr>
<tr><td colspan="2">g</td><td>384</td><td>H0</td></tr>
<tr><td colspan="2">h</td><td>768</td><td>2H0</td></tr>
<tr><td colspan="2">i</td><td>1152</td><td>3H0</td></tr>
<tr><td colspan="2">j</td><td>1536</td><td>3H0</td></tr>
<tr><td colspan="2">k</td><td>1536</td><td>H11</td></tr>
<tr><td colspan="2">l</td><td>1920</td><td>5H0</td></tr>
<tr><td colspan="2">m</td><td>1920</td><td>H12</td></tr>
</table>

<u>Note 1</u> - (Audio coding of mode b3) In addition to H.200/AV.254, higher quality audio coding such as H.200/AV.253 may be used for this mode.

<u>Note 2</u> - For multiple channels of B/H0, all channels are synchronized at the terminal according to § 2.7/H.221.

FIGURE 3.3

multiplex is protected by a double-error correcting code.

- It allows the synchronization of multiple 64 kbit/s or 384 kbit/s connections and the control of the multiplexing of audio, video, data and other signals within the synchronized multiconnection structure in the case of multimedia services such as videoconference.

This Recommendation provides for dynamically subdividing an overall transmission channel of 64 to 1920 kbit/s into lower rates suitable for audio, video, data and telematic purposes. The overall transmission channel is derived by synchronizing and ordering transmissions over from 1 to 6B connections, from 1 to 5 HO connections, or an H11 or H12 connection.

A single 64 kbit/s channel is structured into octets transmitted at 8 kHz.

Bit number								
1	2	3	4	5	6	7	8(SC)	
								1 Octet Number
S	S	S	S	S	S	S	FAS	.
u	u	u	u	u	u	u		8
b	b	b	b	b	b	b		9
-	-	-	-	-	-	-	BAS	.
C	C	C	C	C	C	C		16
h	h	h	h	h	h	h		17
a	a	a	a	a	a	a	(ECS)	.
n	n	n	n	n	n	n		24
n	n	n	n	n	n	n		25
e	e	e	e	e	e	e		.
l	l	l	l	l	l	l		.
#	#	#	#	#	#	#	#	.
1	2	3	4	5	6	7	8	80

FIGURE 3.4 FRAME STRUCTURE OF A SINGLE
64 KBIT/S CHANNEL (B-CHANNEL)

Each bit position of the octets may be regarded as a sub-channel of 8 kbit/s (see Figure 3.4). The eighth sub-channel is called the Service Channel (SC), containing the two critical parts listed below.

- FAS (Frame Alignment Signal): This 8 bit code is used to frame the 80 octets of information in a B channel.

- BAS (Bit-rate Allocation Signal): This 8 bit code describes the capability of a terminal to structure the capacity of the channel or synchronized multiple channels in various ways, and to command a receiver to demultiplex and make use of the constituent signals in such structures. This signal is also used for controls and indications.

Figure 3.5 is a list of the various BAS codes defined in the H.221 Recommendation. The codes are organized into the 8 attributes listed below (columns in Figure 3.5) each of which can have 32 possible values (rows in Figure 3.5).

Attribute	Meaning
000	Audio Coding Command
001	Transfer Rate Command
010	Video and other Command
011	Data Command
100	Terminal Capability 1
101	Terminal Capability 2
110	Reserved
111	Escape Codes

BAS codes provide for the following facilities:

- transmission at various total rates and on single and multiple channels, on clear channels and on networks subject to restrictions to 56 kbit/s and its multiples.
- audio transmission, digitally encoded to various recommended algorithms.
- video transmission, digitally encoded to a recommended algorithm, with provision for future recommended improvement.
- low-speed data (LSD) within the I-channel, or TS1 of a higher rate initial channel.
- high-speed data (HSD) in the highest-numbered 64 kbit/s channel or time-slots (excluding the I-channel).
- data transmission within a multilayer protocol, either in the I-channel (MLP) or in capacity other than the I-channel (H-MLP).
- an encryption control signal

TABLE A1/H.221

BAS numerical values

The column header gives the attribute designation as bits (b_0, b_1, b_2); the left-hand column gives the decimal value of bits $[b_3, b_4, b_5, b_6, b_7]$; for example, "channel #6" has the value (001)[10110]. All unassigned values are reserved, as are values marked (R).

	(000) audio command	(001) trans-fer rate com-mand	(010) other command	(011) LSD/MLP command	(100) audio/ transfer-rate capability	(101) data/ video capabi-lity	(111) escape
[0]	neutral	64	video off	LSD off	neutral	var-LSD	
[1]		2x64	H.261	300	A-law	300	
[2]		3x64	vid-imp(R)	1200	µ-law	1200	
[3]		4x64	video-ISO	4800	G.725-T1	4800	
[4]	A-law, 0U	5x64	AV-ISO	6400	G.725-T2	6400	
[5]	µ-law, 0U	6x64		8000	Au-16kbit/s	8000	
[6]	G.722, m1	384	encryp-on	9600	Au-ISO	9600	
[7]	AU-off, U	2x384	encryp-off	14400		14400	
[8]	Note 1	3x384		16k	128	16k	
[9]	Note 1	4x384		24k	192	24k	
[10]		5x384		32k	256	32k	
[11]		1536		40k		40k	
[12]		1920		48k	512	48k	
[13]	Au-ISO-64	128		56k	768	56k	
[14]	Au-ISO-128	192		62.4k		62.4k	
[15]	Au-ISO-192	256		64k	1152	64k	
[16]	Au-ISO-256		freeze-pic	MLP-off	1B	MLP-4k	HSD
[17]	Au-ISO-384	loss i.c.	fast-update	MLP-4k	2B	MLP-6.4k	H.230
[18]	A-law,0F	chan.#2	Au-loop	MLP-6.4k	3B	var-MLP	Data-apps
[19]	µ-law,0F	chan.#3	Vid-loop	var-MLP	4B		(R-SBE)
[20]		chan.#4	Dig-loop		5B	QCIF	(R-SBE)
[21]		chan.#5	Loop-off	dti-1(R)	6B	CIF	(R-SBE)
[22]		chan.#6		dti-2(R)	restrict	1/29.97	(R-SBE)
[23]		512		dti-3(R)	6B-H0-comp	2/29.97	(R-SBE)
[24]	G.722,m2(Note2)	768			H0	3/29.97	cap-mark
[25]	G.722,m3(Note2)		6B-H0-comp		2H0	4/29.97	start-MBE
[26]	(Au-40k)	1152	No-comp 6B-H0		3H0	V-imp(R)	
[27]	(Au-32k)		restrict		4H0	video-ISO	
[28]	(Au-24k)		derestrict		5H0	AV-ISO	
[29]	Au-16kbit/s	1472			1472	esc-CF(R)	
[30]	(Au-<16k)				H11	encryp.	ns-cap
[31]	Au-off, F			var-LSD	H12	MBE-cap	ns-comm

Note 1 - These codes are listed in Recommendation G.725 with reference to an "application channel"; such a channel has not been defined, the concept having been superseded by that of LSD/MLP; therefore these codes should not be used.

Note 2 - These codes are listed in Recommendation G.725 with reference to "data"; however, the nature of such data (video, LSD, MLP, ECS) must be specified by further commands (001), (010), (011).

FIGURE 3.5

- loopback towards the network for maintenance purposes.
- signalling for control and indications.
- a message system for, inter alia, conveying information concerning equipment manufacturer and type.

In general, there are two types of BAS codes -- 1) command, and 2) capability. When a "command" code is received, the mode of operation is changed. "Capability" BAS codes indicate the ability of a terminal to receive ad properly treat the various types of signals. It follows that having received a set of "capability" codes from the remote terminal Y, terminal X must not transmit signals outside that declared range.

The reader will note in Figure 3.5, that Attribute 111/Value 18 is defined as "Data-apps" which means "applications within LSD/HSD (low speed data/high speed data) channels: a 32-code table, see Table A3/H.221" (Figure 3.6). Figure 3.6 shows the various still picture modes which can be used to supplement the video teleconference.

The reader will also note in Figure 3.5 that Attribute 111/Values 30 and 31 mean "Non-Standard Capability" and "Command" respectively. This permits a H.320 terminal to operate in a proprietary non-standard mode.

3.3 H.242 System for Establishing Communication Between Audiovisual Terminals using Digital Channels up to 2 Mbit/s

Recommendation H.320 defines a number of phases in the establishment of a visual telephone call which precede and succeed the actual communication itself.

- Phase A: Call set-up, out-band signalling
- Phase B1: Mode initialization on initial channel
- Phase CA: Call set-up of additional channel(s), if relevant
- Phase CB1: Initialization on additional channel(s)
- Phase B2 (or CB2): Establishment of common parameters
- Phase C: Visual Telephone Communication
- Phase D: Termination Phase
- Phase E: Call release

Recommendation H.242 defines the detailed "handshake" protocol and

TABLE A3/H.221

Numerical values for applications in LSD/HSD channels

Escape table reached by BAS (111)[18]

The column header gives the attribute designation as bits (b_0, b_1, b_2); the left-hand column gives the decimal value of bits $[b_3, b_4, b_5, b_6, b_7]$. All assigned values are reserved, as are values marked (R).

	capabilities (101)	commands (011)
[0]	ISO-SP baseline on LSD	ISO-SP on in LSD
[1]	ISO-SP baseline on HSD	ISO-SP on in HSD
[2]	ISO-SP spatial	
[3]	ISO-SP progressive	
[4]	ISO-SP arithmetic	
[5]		
[6]		
[7]		
[8]		
[9]		
[10]	Graphics cursor	Cursor data on in LSD
[11]		
[12]		
[13]		
[14]		
[15]		
[16]	Group 3 Fax	Fax on in LSD
[17]	Group 4 Fax	Fax on in HSD
[18]		
[19]		
[20]	V.120 LSD	V.120 LSD
[21]	V.120 HSD	V.120 HSD
[22]		
[23]		
[24]		
[25]		
[26]		
[27]		
[28]		
[29]		
[30]		
[31]		

FIGURE 3.6

procedures which are employed by H.320 terminals in these preliminary phases. Major topics covered in this Recommendation are listed below.

- Basic sequences for in-channel procedures
  - o Capability exchange sequence A
  - o Mode switching sequence B
  - o Frame reinstatement sequence C
- Mode initialization, dynamic mode switching and mode 0 forcing
- Recovery from fault conditions
  - o Unexpected loss of synchronization or frame alignment
  - o Recovery from loss of connection(s)
- Network consideration: Call connection, disconnection and call transfer
- Procedures for activation and deactivation of data channels
- Procedures for operation of terminals in restricted networks
- o 56 Kbit/s network
  - o 56/64 network interworking

3.4 H.230 Frame Synchronous Control and Indication Signals for Audiovisual Systems

Digital audiovisual services are provided by a transmission system in which the relevant signals are multiplexed onto a digital path. In addition to the audio, video, user data and telematic information, these signals include information for the proper functioning of the system. The additional information has been named "control and indication" (C&I) to reflect the fact that while some bits are genuinely for "control", causing a state change somewhere else in the system, others provide for indications to the users as to the functioning of the system.

Recommendation H.230 has two primary elements. First, it defines the C&I symbols related to video, audio, maintenance, and multipoint. Second, it contains a table of BAS escape codes which clarifies the circumstances under which some C&I functions are mandatory and others optional. The reader will note on Figure 3.5 that the H.230 escape code is signalled by Attribute 111 and Value 17.

3.5 Audio Coding (AV.250)

The BAS codes of H.221 are used to signal a wide range of possible audio coding modes. The most prominent modes define existing CCITT

Recommendations G.711 and G.722. Recommendation G.711 (Pulse Code Modulation of Voice Frequencies) is used for narrowband speech since it samples only at 8,000 samples/sec. and encodes to 8 bits/sample for a transmission rate of 64 Kbps. Recommendation G.722 (7kHz Audio-Coding with 64 Kbit/s) describes the characteristics of an audio (50 to 7 000 Hz) coding system which may be used for a variety of higher quality speech applications. The coding system uses sub-band adaptive differential pulse code modulation (SB-ADPCM) within a bit rate of 64 Kbit/s. In the SB-ADPCM technique used, the frequency band is split into two sub-bands (higher and lower) and the signals in each sub-band are encoded using ADPCM. The system has three basic modes of operation corresponding to the bit rates used for 7 kHz audio coding: 64, 56, and 48 kbit/s.

BAS codes have been reserved for future audio standards which are now in development. Bit rates include 40, 43, 24, and 16 Kbps. Of particular importance is the 16 Kbps (AV.254) which is necessary to provide a complete audio visual service in one B channel.

3.6 Multipoint

At this time, standards do not exist for multipoint operation of H.320/P x 64 terminals. However, work is underway on three CCITT Recommendations to fill this void. Copies of these three Recommendations illustrating their very preliminary form are included in Appendix C.

AV.231	Multipoint control unit for audiovisual services
AV.243	System for establishing communication between three or more audiovisual terminals using digital channels up to 2mbit/s
AV.440	Call-control procedures for real-time audiovisual conference calls

4.0 OVERVIEW OF RECOMMENDATION H.261 (VIDEO CODEC)

Figure 4.1 is a functional block diagram of the video codec as defined in Recommendation H.261. The heart of the system is the source coder which compresses the incoming video signal by reducing redundancy inherent in the TV signal. The multiplexer combines the compressed data with various side information which indicates alternative modes of operation. A transmission buffer is employed to smooth the varying bit rate from the source encoder to adapt it for the fixed bit rate communication channel. A transmission coder includes functions such as forward error control to prepare the signal for the data link.

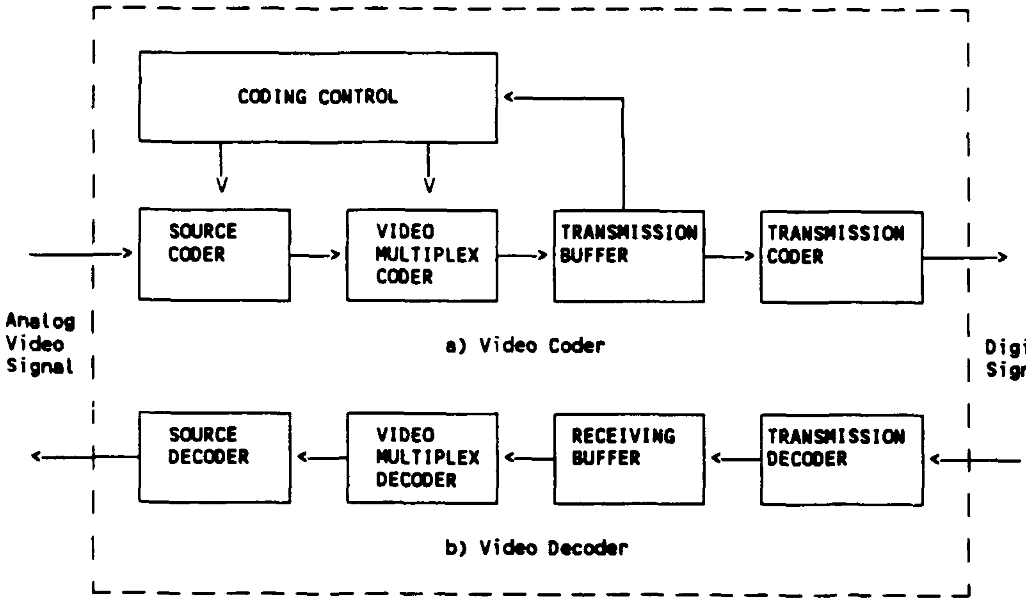


FIGURE 4.1 BLOCK DIAGRAM OF THE VIDEO CODEC

One of the most challenging problems to be solved by the codec was the reconciliation of the incompatibility between European TV standards (PAL, SECAM) and those in most other areas of the world (NTSC). PAL and SECAM employ 625 lines and a 50 Hz field rate while NTSC has 525 lines and a 60 Hz field rate. This conflict was resolved by adopting a Common Intermediate Format (CIF) and QCIF (Quarter CIF) as the picture structure which must be employed for any transmission adhering to H.261. The CIF and QCIF parameters are defined in Table 4.1.

The QCIF format, which employs half the CIF spatial resolution in both horizontal and vertical directions, is the mandatory H.261 format: full CIF is

	CIF	QCIF
Coded Pictures per Second	29.97	(or integral submultiples)
Coded Luminance pixels per line	352	176
Coded Luminance lines per picture	288	144
Coded Color pixels per line	176	88
Coded Color lines per picture	144	72

TABLE 4.1 CIF AND QCIF PARAMETERS

optional. It is anticipated that QCIF will be used for videophone applications where head-and-shoulders pictures are sent from desk to desk. Conversely, it is assumed that the full CIF format will be used for teleconferencing where several people must be viewed in a conference room.

Figure 4.2 is a functional block diagram outlining the H.261 source coder. Interframe prediction is first carried out in the pixel domain. The prediction errors are encoded by the Discrete Cosine Transform using blocks of 8 pels x 8 pels. The Transform coefficients are next quantized and fed to the multiplexer. Motion compensation is included in the prediction on an optional basis.

PICTURE STRUCTURE

In the encoding process, each picture is subdivided into Groups of Blocks (GOB). As shown in Figure 4.3, the CIF picture is divided into 12 GOB's while QCIF has only three GOB's. From the GOB level down, the structure of CIF and QCIF is identical. A header at the beginning of the GOB permits resynchronization and changing the coding accuracy.

Each GOB is further divided into 33 macroblocks, as shown in Figure 4.4. The macroblock header defines the location of the macroblock within the GOB, the type of coding to be performed, possible motion vectors, and which blocks within the macroblock will actually be coded. There are two basic types of coding. In Intra coding, coding is performed without reference to previous pictures. This mode is relatively rare, but is required for forced updating, and every macroblock

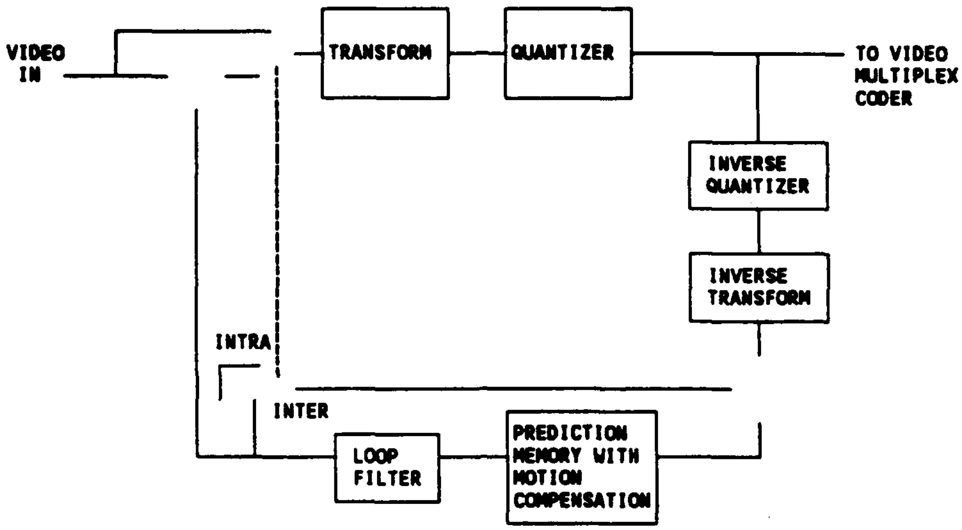


FIGURE 4.2 SOURCE CODER

1	2
3	4
5	6
7	8
9	10
11	12

CIF

1
3
5

QCIF

FIGURE 4.3 ARRANGEMENT OF GOBs IN A PICTURE

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

FIGURE 4.4 ARRANGEMENT OF MACROBLOCKS IN A GOB

must occasionally be Intra coded to control the accumulation of inverse transform mismatch error. The more common coding type is Inter, in which only the difference between the previous picture and the current one is coded. Of course, for picture areas without motion, the macroblock does not have to be coded at all.

Each macroblock is further divided into six blocks, as shown in Figure 4.5.

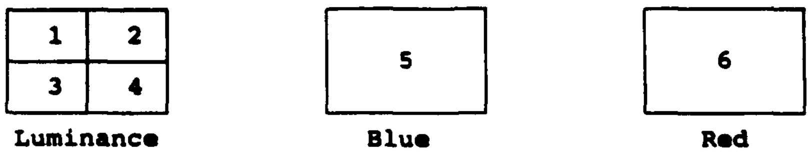


FIGURE 4.5 ARRANGEMENT OF BLOCKS IN A MACROBLOCK

Four of the blocks represent the luminance, or brightness, while the other two represent the red and blue color differences. Each block is 8 x 8 pixels, so it can be seen that the color resolution is half of the luminance resolution in both dimensions.

<u>EXAMPLE OF BLOCK CODING</u>

Figure 4.6 shows a simple example of how each 8 x 8 block is coded. In this case, Intra coding is used, but the principle is the same for Inter coding. Figure 4.6a shows the original block to be coded. Without compression, this would take 8 bits to code each of the 64 pixels, or a total of 512 bits. First, the block is transformed, using the two-dimensional Discrete Cosine Transform (DCT), giving the coefficients of Figure 4.6b. Note that most of the energy is concentrated into the upper left-hand corner of the coefficient matrix. Next, the coefficients of Figure 4.6b are quantized with a step size of 6. (The first term {DC} always uses a step size of 8.) This produces the values of Figure 4.6c, which are much smaller in magnitude than the original coefficients and most of the coefficients become zero. The larger the step size, the smaller the values produced, resulting in more compression.

The coefficients are then reordered, using the Zig-Zag scanning order of Figure 4.7. All zero coefficients are replaced with a count of the number of zero's before each non-zero coefficient (RUN). Each combination of RUN and VALUE produces a Variable Length Code (VLC) that is sent to the decoder. The last non-zero VALUE is followed by an End of Block (EOB) code. The total number of bits used to describe the block is 25, a compression of 20:1.

At the decoder (and at the coder to produce the prediction picture), the step size and VALUE's are used to reconstruct the inverse quantized coefficients,

75	76	77	78	79	80	81	82
77	78	79	80	81	82	83	84
79	80	81	82	83	84	85	86
81	82	83	84	85	86	87	88
83	84	85	86	87	88	89	90
85	86	87	88	89	90	91	92
87	88	89	90	91	92	93	94
89	90	91	92	93	94	95	96

a) ORIGINAL BLOCK (8x8x8 = 512 BITS)

76	76	77	79	80	81	82	83
77	77	78	80	81	82	83	84
79	79	80	81	83	84	85	86
81	82	83	84	85	87	88	88
84	84	85	87	88	89	90	91
86	87	88	89	91	92	93	93
88	89	90	91	92	94	95	95
89	90	91	92	93	95	96	96

f) RECONSTITUTED BLOCK

684	-19	-1	-2	0	-1	0	-1
-37	0	-1	0	0	0	0	-1
0	0	0	0	0	0	0	0
-4	-1	-1	-1	-1	0	-1	-1
0	0	0	0	0	0	0	0
-2	0	0	-1	0	-1	0	-1
0	0	0	0	-1	-1	-1	-1
-1	-1	-1	0	-1	0	-1	0

b) TRANSFORMED BLOCK COEFFICIENTS

688	-21	0	0	0	0	0	0
-39	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

e) INVERSE QUANTIZED COEFFICIENTS

86	-3	0	0	0	0	0	0
-6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

c) QUANTIZED COEFFICIENT LEVELS

RUN	LEVEL	CODE
0	86	01010110
0	-3	001011
0	-6	001000011
	EOB	10

TOTAL CODE LENGTH = 25

d) COEFFICIENTS IN ZIG-ZAG ORDER AND VARIABLE LENGTH CODED

FIGURE 4.6 SAMPLE INTRA BLOCK CODING

which, as shown in Figure 4.6e are similar to, but not exactly equal to, the original coefficients. When these coefficients are inverse transformed, the result of Figure 4.6f is obtained. Note that the differences between this block and the original block are quite small.

MOTION COMPENSATION

The operation of motion compensation is shown in Figure 4.8. Block "A" is a block in the current picture that is to be coded. Block "B" is the block at the same position as "A" but in the picture that was previously stored in both coder and decoder. Because of image motion, block "A" more closely resembles the pixel data from block "C" than that from block "B". The displacement of block "C" from block "B", measured in pixels in x and y directions, is the motion vector. The pixel-by-pixel difference between blocks "A" and "C" is transformed and coded. The motion vector and code data are transmitted to the decoder, where the inverse transformed block data is added to the data in block "C" pointed to by the motion

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

FIGURE 4.7 SCANNING ORDER IN A BLOCK

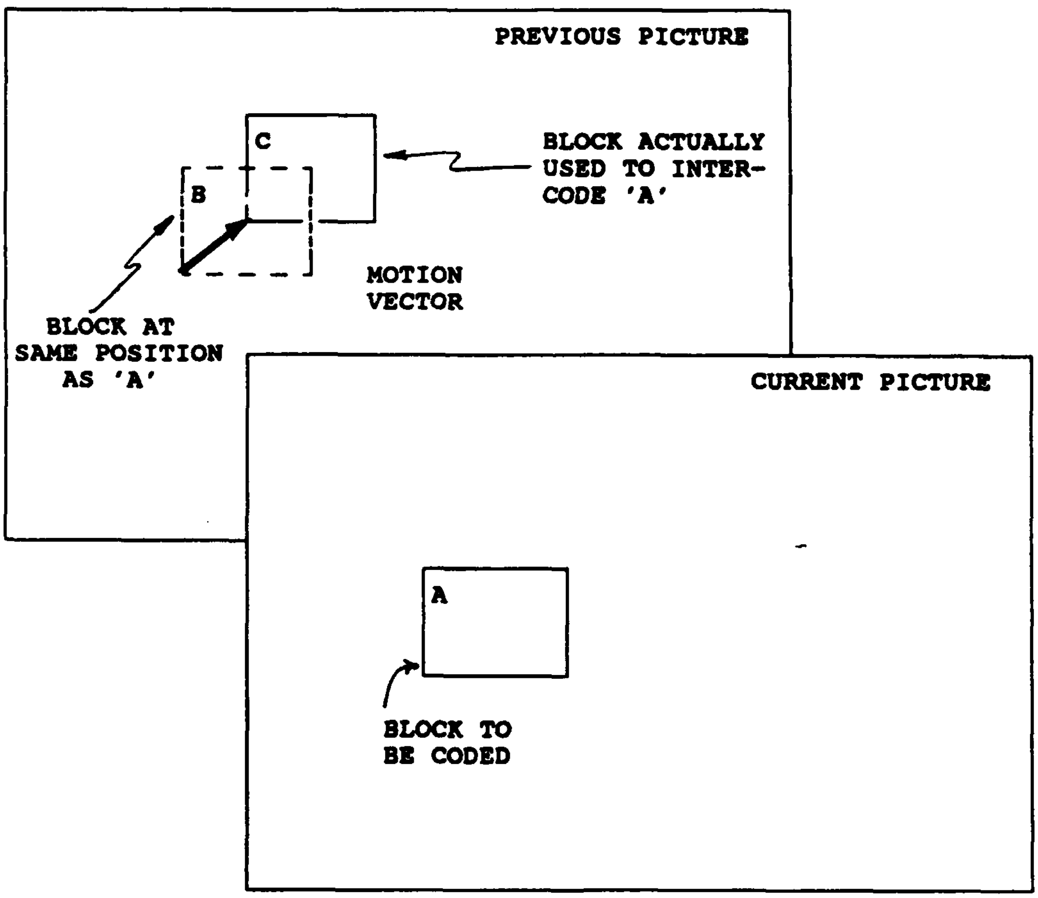


FIGURE 4.8 INTER-FRAME CODING WITH MOTION VECTORS

vector, and placed in the block "A" position.

The use of motion vectors is optional in the coder, where the calculation of the optimum motion vectors is complex, but required in the decoder, where the reconstruction of the motion is relatively simple.

FLEXIBILITY OF STANDARD

The H.261 standard does not define all aspects of image coding and decoding. Rather it is just a interoperability specification, guaranteeing that any codecs manufactured according to the standard will be able to communicate with each other. This still allows considerable freedom for manufacturers to offer better performance, and new developments may be able to be incorporated. (This is in contrast with the G.722 audio standard, where the coding algorithm is rather precisely defined.) For example, the encoder strategy is not defined. Which blocks will be encoded, with what type of code, and with what accuracy is under control of the designer. While there is less freedom for the decoder, post processing, such as filtering or interpolation, can be used to improve the appearance of the image and is under control of the designer.

Furthermore, the P x 64 standards permit two codecs to negotiate to an altogether different algorithm that they both incorporate, with the H.261 algorithm becoming a fall-back mode for codecs of different manufacture.

5.0 SUMMARY AND CONCLUSIONS

Work during this report interval on the P x 64 standards has been unusually productive. Five CCITT Recommendations (H.320, H.261, H.221, H.242, H.230) were finalized in December, 1990. These Recommendations completely define an audio visual terminal for videophone and video teleconferencing applications. Products conforming to these Recommendations are already in the marketplace, and it is anticipated that a revolution in video telephony will occur which will be similar to that which occurred in the facsimile market when the Group 3 standard was finalized. This revolution will occur for the following reasons.

- The standard provides picture quality which is at least competitive with proprietary systems in the marketplace.
- The picture quality will continually improve due to competitive market pressure.
- The system is very flexible permitting a wide range of operational modes -- video operation, voice quality, still imagery, transmission bit rates, etc.
- Codec cost will rapidly drop due to the availability of low cost VLSI chips and high volume production.
- Communication costs will drop as the ISDN is introduced and traffic volume increases.

APPENDIX A

P x 64 RECOMMENDATIONS OF THE H-SERIES

International Telegraph and Telephone
Consultative Committee
(CCITT)

Period 1989-1992

COM XV-R 37-E
August 1990
Original: English

Questions: 3, 4/XV

STUDY GROUP XV - REPORT R 37

SOURCE: STUDY GROUP XV (GENEVA MEETING, 16-27 JULY 1990)

TITLE: RECOMMENDATIONS OF THE H-SERIES

CONTENTS

CCITT\COMXV\RAPP\R037E1.TXS

1. Recommendation H.221

FRAME STRUCTURE FOR A 64 TO 1920 kbit/s CHANNEL IN AUDIOVISUAL TELESERVICES[1]

CONTENTS

[1] This Recommendation completely replaces the text of Recommendations H.221 and H.222 published in Fascicle III.6 of the Blue Book.

Introduction

The purpose of this Recommendation is to define a frame structure for audiovisual teleservices in single or multiple B or H0 channels or a single H11 or H12 channel which makes the best use of the characteristics and properties of the audio and video encoding algorithms, of the transmission frame structure and of the existing CCITT Recommendations. It offers several advantages:

- It takes into account CCITT Recommendations such as G.704, X.30/I.461, etc. It may allow the use of existing hardware or software.

- It is simple, economic and flexible. It may be implemented on a simple microprocessor using well-known hardware principles.

- It is a synchronous procedure. The exact time of a configuration change is the same in the transmitter and the receiver. Configurations can be changed at 20 ms intervals.

- It needs no return link for audiovisual signal transmission, since a configuration is signalled by repeatedly transmitted codewords.

- It is very secure in case of transmission errors, since the code controlling the multiplex is protected by a double-error correcting code.

- It allows the synchronization of multiple 64 kbit/s or 384 kbit/s connections and the control of the multiplexing of audio, video, data and other signals within the synchronized multiconnection structure in the case of multimedia services such as videoconference.

- It can be used to derive octet synchronization in networks where this is not provided by other means.

- It can be used in multipoint configurations, where no dialogue is needed to negotiate the use of a data channel.

- It provides a variety of data bit-rates (from 300 bit/s up to almost 2 Mbit/s) to the user.

1. Basic principle

This Recommendation provides for dynamically subdividing an overall transmission channel of 64 to 1920 kbit/s into lower rates suitable for audio, video, data and telematics purposes. The overall transmission channel is derived by synchronizing and ordering transmissions over from 1 to 6 B connections, from 1 to 5 H0 connections, or an H11 or H12 connection. The first connection established is the initial connection and carries the <u>initial channel</u> in each direction. The additional connections carry <u>additional channels</u>.

The total rate of transmitted information is called the "transfer rate"; it is possible to fix the transfer rate less than the capacity of the overall transmission channel (values listed in Annex 1).

A single 64 kbit/s channel is structured into octets transmitted at 8 kHz. Each bit position of the octets may be regarded as a sub-channel of 8 kbit/s (see Figure 1a). The eighth sub-channel is called the Service Channel (SC), consisting of several parts as described in sections 1.1 - 1.4 below.

An H0, H11 or H12 channel may be regarded as consisting of a number of 64 kbit/s <u>time-slots</u> (TS) (see Figure 1b). The lowest numbered time-slot is structured exactly as described for a single 64 kbit/s channel, while the other TS have no such structure. In the case of multiple B or H0 channels, all channels have a frame structure; that in the initial channel controls most functions across the overall transmission, while the frame structure in the additional channels is used for synchronization, channel numbering and related controls.

The term "I-channel" is applied to the initial or only B-channel, to TS1 of initial or only H0 channel, and to TS1 of H11, H12 channels.

Bit number 1	2	3	4	5	6	7	8(SC)	
								1 Octet Number
S	S	S	S	S	S	S	FAS	.
u	u	u	u	u	u	u		8
b	b	b	b	b	b	b		9
-	-	-	-	-	-	-	BAS	.
C	C	C	C	C	C	C		16
h	h	h	h	h	h	h		17
a	a	a	a	a	a	a	(ECS)	.
n	n	n	n	n	n	n		24
n	n	n	n	n	n	n		25
e	e	e	e	e	e	e		.
l	l	l	l	l	l	l		.
#	#	#	#	#	#	#	#	.
1	2	3	4	5	6	7	8	80

FIGURE 1(a)/H.221

<u>Frame structure of a single 64 kbit/s channel (B-channel)</u>

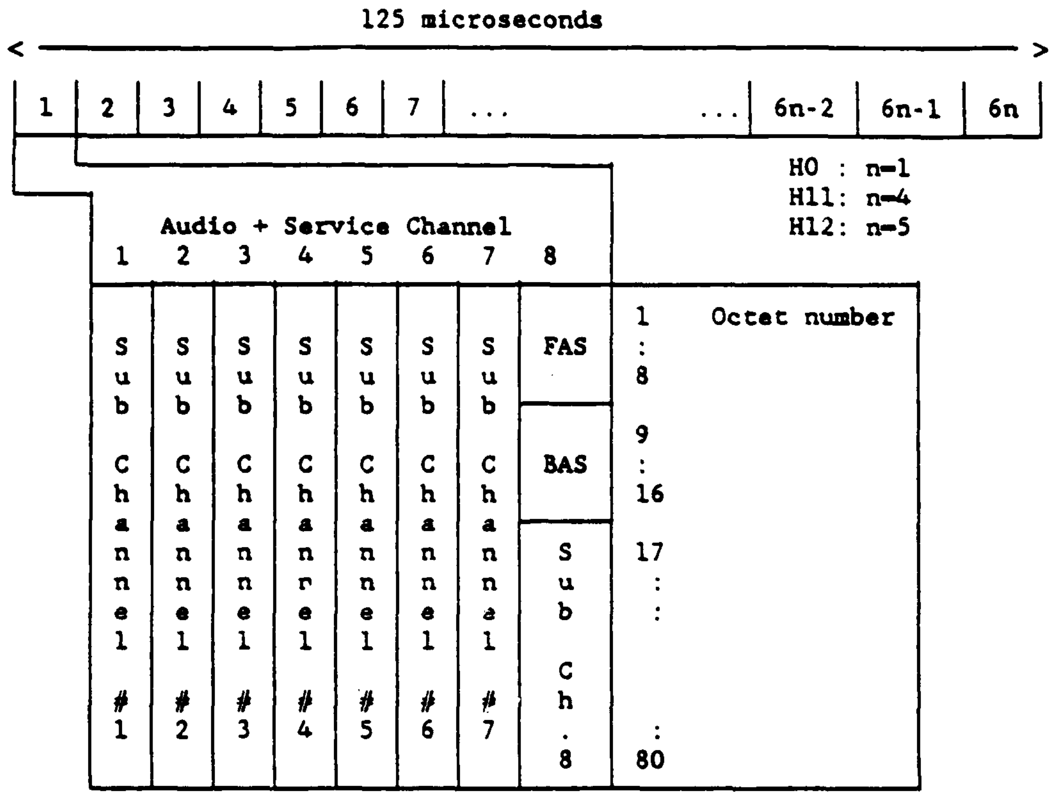


FIGURE 1(b)/H.221

Frame structure of higher-rate single channels
(H0, H11, H12 channels)

1.1 Frame alignment signal (FAS)

This signal structures the I-channel and other framed 64 kbit/s channels into frames of 80 octets each and multiframes (MF) of 16 frames each. Each multiframe is divided into eight 2-frame sub-multiframes (SMF). The term "frame alignment signal" (FAS) refers to bits 1-8 of the SC in each frame. In addition to framing and multiframing information, control and alarm information may be inserted in the FAS, as well as error check information to control end-to-end error performance and to check frame alignment validity. Other time-slots are aligned to the first.

The bits are transmitted to line in order, bit 1 first.

When an 8 kHz network clock is provided, FAS is transmitted and received in the least significant bit of the octet within each 125 microsecond, e.g., in an ISDN basic or primary rate interface.

The FAS can be used to derive receive octet timing when it is not provided by the network. However, in the latter case, the terminal cannot transmit FAS with correct alignment into the octet timed part of the network and cannot intercommunicate with terminals which rely only on network timing for octet alignment.

1.2 Bit-rate allocation signal (BAS)

Bits 9-16 of the SC in each frame are referred to as BAS. This signal allows the transmission of codewords to describe the capability of a terminal to structure the capacity of the channel or synchronized multiple channels in various ways, and to command a receiver to demultiplex and make use of the constituent signals in such structures. This signal is also used for controls and indications.

Note - For some countries having 56 kbit/s channels, the net available bit rates will be 8 kbit/s less. Interworking between a 64 kbit/s terminal and a 56 kbit/s terminal is established according to the frame structure in Annex 2.

1.3 Encryption control signal (ECS)

A future encryption capability may require a dedicated transmission channel. It is anticipated that 800 bit/s should be provided when required by allocating the bits 17-24 of the Service Channel. This reduces variable data and video transmission rates herein by 800 bit/s. The 800 bit/s is referred to as the ECS Channel.

1.4 Remaining capacity

The remaining capacity (including the rest of the service channel), carried in bits 1-8 of each octet in the case of a single 64 kbit/s connection, may convey a variety of signals within the framework of a multimedia service, under the control of the BAS. Some examples follow:

- voice encoded at 56 kbit/s using a truncated form of PCM of CCITT Recommendation G.711 (A-law or μ-law);
- voice encoded at 16 kbit/s and video at 46.4 kbit/s;
- voice encoded at 56 kbit/s with a bandwidth 50 - 7000 Hz (subband ADPCM according to CCITT Recommendation G.722); the coding algorithm is also able to work at 48 kbit/s - data can then be dynamically inserted at up to 14.4 kbit/s;
- still pictures coded at 56 kbit/s;
- data at 56 kbit/s inside an audiovisual session (e.g., file transfer for communicating between personal computers).

2. Frame alignment

2.1 General

An 80-octet frame length produces an 80-bit word in the Service Channel. These 80 bits are numbered 1-80. Bits 1-8 of the Service Channel in every frame constitute the FAS (see Figure 2/H.221), whose content is as follows:

- multiframe structure (see section 2.2);

- frame alignment word (FAW);
- "A-bit";
- "E-" and "C-bits" (see section 2.6).

SUCCESSIVE FRAMES	Bit # 1	2	3	4	5	6	7	8
Even frames	Note 1	0	0	1	1	0	1	1
			Frame Alignment Word - Note 2					
Odd frames		1	A	E	C1	C2	C3	C4
	Note 1	Note 2	Note 3	Note 4				

Note 1 - See section 2.2 and Figure 3/H.221.

Note 2 - The first seven bits of the Frame Alignment Word are in the even frames. The eighth bit of the FAW in the odd frame is the complement of the first FAW bit in order to avoid simulation of FAW by a frame-repetitive pattern.

Note 3 - A-bit loss of multiframe alignment indication (0 = alignment; 1 = loss).

Note 4 - The use of bits E and C1-C4 is described in section 2.6 (0 = no error or CRC not in use; 1 = error).

FIGURE 2/H.221

Assignment of bits 1-8 of the service channel in each frame

The FAW consists of "0011011" in bits 2-8 of the FAS in even frames, complemented by an "1" in bit 2 of the succeeding odd frame.

The "A-bit" of the I-channel is set to zero whenever the receiver is in multiframe alignment, and is set to "1" otherwise (see section 2.3); for additional channels, see section 2.7.1.

2.2 Multiframe structure

Each multiframe contains 16 consecutive frames numbered 0 to 15 divided into eight sub-multiframes of two frames each (Figure 3). The multiframe alignment signal is located in bit 1 of frames 1-3-5-7-9-11 and has the form 001011. Bit 1 of frame 15 remains reserved for future use. The value is fixed at 0.

Bit 1 of frames 0-2-4-6 may be used for a modulo 16 counter to number multiframes in descending order. The least significant bit is transmitted in frame 0, and the most significant bit in frame 6. The receiver uses the multiframe numbering to equalize out the differential delay of separate connections, and to synchronize the received signals.

Bit 1 of frame 8 is set to 1 when multiframes are numbered and is set to 0 when they are not.

Bit 1 of frames 10-12-13 must be used to number each channel in a multiconnection structure so that the distant receiver can place the octets received in each 125 microseconds in the correct order.

Information bits in the multiframe should be validated by, for example, being received consistently for three multiframes.

	Sub-multi-frame	Frame	Bits 1 to 8 of the Service Channel in every frame							
			1	2	3	4	5	6	7	8
Multiframe		0	N1	0	0	1	1	0	1	1
	SMF 1	1	0	1	A	E	C1	C2	C3	C4
		2	N2	0	0	1	1	0	1	1
	SMF 2	3	0	1	A	E	C1	C2	C3	C4
		4	N3	0	0	1	1	0	1	1
	SMF 3	5	1	1	A	E	C1	C2	C3	C4
		6	N4	0	0	1	1	0	1	1
	SMF 4	7	0	1	A	E	C1	C2	C3	C4
		8	N5	0	0	1	1	0	1	1
	SMF 5	9	1	1	A	E	C1	C2	C3	C4
		10	L1	0	0	1	1	0	1	1
	SMF 6	11	1	1	A	E	C1	C2	C3	C4
		12	L2	0	0	1	1	0	1	1
	SMF 7	13	L3	1	A	E	C1	C2	C3	C4
		14	TEA	0	0	1	1	0	1	1
	SMF 8	15	R	1	A	E	C1	C2	C3	C4

L1-L3: channel number, least significant bit in L1

Channel	L3	L2	L1
Initial	0	0	1
Second	0	1	0
Third	0	1	1
...	..	..	..
Sixth	1	1	0

R: Reserved for future use - set to 0.
A, E, C1-C4: As in Figure 2/H.221.
N1-N4: Used for multiframe numbering as described in section 2.2. Set to 0 while numbering is inactive.

<u>Multiframe number</u>:	N4	N3	N2	N1	
0	0	0	0	0	(or numbering inactive)
1	0	0	0	1	
2	0	0	1	0	
...	...	...	...	...	
15	1	1	1	1	

N5: Indicates whether multiframe numbering is active (N5=1) or inactive (N5=0).
TEA: The terminal equipment alarm is set to 1 in the outgoing signal while an internal terminal equipment fault exists such that it cannot receive and act on the incoming signal. Otherwise it is set to 0.

FIGURE 3/H.221

<u>Assignment of bits 1-8 of the service channel in each frame in a multiframe</u>

2.3 Loss and recovery of frame alignment

Frame alignment is defined to have been lost when three consecutive frame alignment words nave been received with an error.

Frame alignment is defined to have been recovered when the following sequence is detected:

- for the first time, the presence of the correct first seven bits of the frame alignment word;
- the eighth bit of the frame alignment word in the following frame is detected by verifying that bit 2 is a 1;
- for the second time, the presence of the correct first seven bits of the frame alignment word in the next frame.

If frame alignment is achieved but multiframe alignment cannot be achieved, then frame alignment should be sought at another position.

When the frame alignment is lost, A-bit of the next odd frame is set to 1 in the transmit direction.

2.4 Loss and recovery of multiframe alignment

Multiframe alignment is needed to number and synchronize two or more channels, and possibly also for encryption. Terminals such as those having only single-channel capabilities which have no use for the multiframe structure must transmit the multiframe structure, but need not check for multiframe alignment on the incoming signal: they may transmit outgoing A=0 when frame alignment is recovered. (Note - such a terminal cannot transmit TEA - see Figure 3/H.221).

After multiframe alignment has been validated the other functions represented by bit 1 of the Service Channel can be used. When multiframe alignment of the distant terminal has been signalled (A=0 received) the distant terminal is expected to have validated BAS codes and to be able to interpret BAS codes.

Multiframe alignment is defined to have been lost when three consecutive multiframe alignment signals have been received with an error. It is defined to have been recovered when the multiframe alignment signal has been received with no error in the next multiframe. When multiframe alignment is lost, even when an unframed mode is received, the A-bit of the next odd frame is set to 1 in the transmit direction. It is reset to 0 when multiframe alignment is regained. It is reset in additional channels when multiframe alignment and synchronism with the initial channel is regained.

2.5 Procedure to recover octet timing from frame alignment

When the network does not provide octet timing, the terminal may recover octet timing in the receive direction from bit timing and from the frame alignment. The octet timing in the transmit direction may be derived from the network bit timing and an internal octet timing.

2.5.1 General rule

The receive octet timing is normally determined from the FAS position. But at the start of the call and before the frame alignment is gained, the receive octet timing may be taken to be the same as the internal transmit octet timing. As soon as a first frame alignment is gained, the receive octet timing is initialized at the new bit position, but it is not yet validated. It will be validated only when frame alignment is not lost during the next 16 frames.

2.5.2 Particular cases

a) When, at the initiation of a call, the terminal is in a forced reception mode, or when the frame alignment has not yet been gained, the terminal may temporarily use the transmit octet timing.

b) When frame alignment is lost after being gained, the receive octet timing should not change until frame alignment is recovered.

c) As soon as frame and multiframe alignment have been gained once, the octet timing is considered as valid for the rest of the call, unless frame alignment is lost and a new frame alignment is gained at another bit position.

d) When the terminal switches from a framed mode to an unframed mode (by means of the BAS), the octet timing previously gained must be kept.

e) When a new frame alignment is gained on a new position, different from that previously validated, the receive octet timing is reinitialized to the new position but not yet validated and the previous bit position is stored. If no loss of frame alignment occurs in the next 16 frames, the new position is validated, otherwise the stored old bit position is reutilized.

2.5.3 Search for frame alignment signal (FAS)

Two methods may be used: sequential or parallel. In the sequential method, each of the eight possible bit positions for the FAS is tried. When FAS is lost after being validated, the search must resume starting from the previously validated bit position. In the parallel method, a sliding window, shifting one bit for each bit period, may be used. In that case, when frame alignment is lost, the search must resume starting from the bit position next to the previously validated one.

2.6 Description of the CRC4 procedure

In order to provide an end-to-end quality monitoring of the connection, a 4-bit cyclic redundancy check (CRC4) procedure may be used and the four bits C1, C2, C3 and C4 computed at the source location are inserted in bit positions 5 to 8 of the odd frames. In addition, bit 4 of the odd frames, the "E-bit", is used to transmit an indication as to whether the most recent CRC block, received in the incoming direction, contained errors or not.

When the CRC4 procedure is not used, bit E shall be set to 0, and bits C1, C2, C3 and C4 shall be set to 1 by the transmitter. Provisionally, the receiver may disable reporting of CRC errors after receiving eight consecutive CRCs set to all 1s, and it may enable reporting of CRC errors after receiving two consecutive CRCs each containing a 0 bit.

2.6.1 Computation of the CRC4 bits

The CRC4 bits C1, C2 C3 and C4 are computed for each B/H0/H11/H12 channel[1], for a block made of two frames: one even frame (containing the first seven bits of FAW) followed by one odd frame (containing the eighth bit of FAW). The CRC4 block size is then 160/960/3840/4800 octets for a B/H0/H11/H12 channel[1] and the computation is performed 50 times per second.

Note - This is still valid for the case of H0/H11 in restricted networks, the stuffed bits being included in the computation. For restricted B, see Annex 2.

2.6.1.1 Multiplication-division process

A given C1-C4 word located in block N is the remainder after multiplication by x^4 and then division (modulo 2) by the generator polynomial $x^4 + x + 1$ of the polynomial representation of block (N-1).

When representing contents of a block as a polynomial, the first bit in the block should be taken as being the most significant bit. Similarly C1 is defined to be the most significant bit of the remainder and C4 the least significant bit of the remainder.

This process can be realized with a four-stage register and two exclusive-ORs.

2.6.2.1 Encoding procedure

i) The CRC bit positions in the odd frame are initially set at zero, i.e., C1=C2=C3=C4=0.

ii) The block is the acted upon by the multiplication-division process referred to above in 2.6.1.1.

iii) The remainder resulting from the multiplication-division process is stored ready for insertion into the respective CRC locations of the next odd frame.

Note - These CRC bits do not affect the computation of the CRC bits of the next block, since the corresponding locations are set at zero before the computation.

2.6.1.3 Decoding procedure

i) A received block is acted upon by the multiplication-division process, referred to above in 2.6.1.1, after having its CRC bits extracted and replaced by zeros.

---

[1] If the transfer rate is such that a part of any H0/H11/H12 channel is unoccupied, then the computation is made only for that part covered by the transfer rate.

ii) The remainder resulting from this multiplication-division process is then stored and subsequently compared on a bit-by-bit basis with the CRC bits received in the next block.

iii) If the decoded calculated remainder exactly corresponds to the CRC bits sent from the encoder, it is assumed that the checked block is error-free.

2.6.2 Consequent actions

2.6.2.1 Action on bit E

Bit E of block N is set to 1 in the transmitting direction if bits C1-C4 detected in the most recent block in the opposite direction have been found in error (at least one bit in error). In the opposite case it is set to zero.

2.6.2.2 Monitoring for incorrect frame alignment (see note below)

In the case of a long simulation of the FAW, the CRC4 information can be used to re-invite a search for frame alignment. For such a purpose it is possible to count the number of CRC blocks in error within two seconds (100 blocks) and to compare this number with 89. If the number of CRC blocks in error is greater than or equal to 89, a search for frame alignment should be re-initiated.

These values 100 and 89 have been chosen in order that:

- For a random transmission error rate of 10^{-3}, the probability of incorrectly re-initiating a search for frame alignment, because of 89 or more blocks in error, should be less than 10^{-4}.

- In case of simulation of frame alignment, the probability of not re-initiating a search of frame alignment after a two-second period should be less than 2.5%.

Note - Values in this and the next section exemplify the case of a 64 kbit/s channel. For H0, H11 or H12 channels the details will differ but the principles are still applicable.

2.6.2.3 Monitoring for error performance

The quality of the 64 kbit/s connection can be monitored by counting the number of CRC blocks in error within a period of one second (50 blocks). For instance, a good evaluation of the proportion of seconds without errors as defined in CCITT Recommendation G.821 can be provided.

For information purposes, the following proportions of CRC block in error can be computed for randomly distributed errors of error rate P_e:

P_e	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Percentage of CRC blocks in error	70%	12%	1.2%	0.12%	0.012%

By counting the received E bits, it is possible to monitor the quality of the connection in the opposite direction.

2.7 Synchronization of multiple connections

Some audiovisual terminals will be able to communicate over multiple B or H0 connections (see note below). In this case, a single B or H0 initial connection is established, the possibility for more connections is determined from the transfer rate capability BAS of Annex 1 and the additional connections are then established and synchronized by the terminal using the multiframe structure.

Note - A connection is an individual call between the terminals. A channel is the transmission in one direction over the connection.

2.7.1 Multiple B connections

FAS and BAS are transmitted in each B-channel.

FAS operation is as follows:

- multiframe numbering is used to determine relative transmission delay between B-channels as described in 2.2;
- the channel numbers are transmitted as described in 2.2 with the channel of the initial connection being numbered 1 and there being up to five additional connections;
- the outgoing A-bit is set to 1 in the additional B-channel of the same connection whenever the received additional channel is not synchronized to the initial channel;
- when receive synchronization is achieved between the initial and additional channels by introducing delay to align their respective multiframe signals, the transmitted A-bit is set to 0;
- the E-bit for each additional B-channel is transmitted in the additional B-channel in the same connection, because it relates to a physical condition of the transmission path.

The BAS operation in additional connections is restricted to the transmission of the additional channel number (thus the channel numbering must be sent both in BAS according to Annex 1 and in the FAS as in section 2.2).

The distant terminal, upon receiving the A-bit set to zero with respect to sequentially numbered channels, can add their capacity to the initial connection by sending the transfer rate BAS in Annex 1. The order of the bits transmitted in the channels is in accordance with the examples given in Figure 4/H.221.

2.7.2 Multiple H0 connections

FAS and BAS are transmitted in the first time-slot of each H0.

FAS operation is as in 2.7.1 except that the channel number is used to order the six octets received each 125 microseconds with respect to the six octet groups received in other channels.

The BAS operation in additional channels is as specified in 2.7.1.

3. Bit-rate allocation signal

3.1 Encoding of the BAS

The bit-rate allocation signal (BAS) occupies bits 9-16 of the Service Channel in every frame. An eight bit BAS code (b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7) is complemented by eight error correction bits (p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7) to implement a (16,8) double error correcting code. This error correcting code is obtained by shortening the (17,9) cyclic code with generator polynomial:

$$g(x) = x^8 + x^7 + x^6 + x^4 + x^2 + x + 1$$

The error correction bits are calculated as coefficients of the remainder polynomial in the following equation:

$$p_0x^7 + p_1x^6 + p_2x^5 + p_3x^4 + p_4x^3 + p_5x^2 + p_6x + p_7$$

$$= RES_{g(x)}[b_0x^{15} + b_1x^{14} + b_2x^{13} + b_3x^{12} + b_4x^{11} + b_5x^{10} + b_6x^9 + b_7x^8]$$

where $RES_{g(x)}[f(x)]$ represents the residue obtained by dividing $f(x)$ by $g(x)$.

The BAS code is sent in the even-numbered frame, while the associated error correction bits are sent in the subsequent odd-numbered frame. The bits of the BAS code or the error correction are transmitted in the following order to avoid emulation of the frame alignment word:

Bit position	Even frame	Odd frame
9	b_0	p_2
10	b_3	p_1
11	b_2	p_0
12	b_1	p_4
13	b_5	p_3
14	b_4	p_5
15	b_6	p_6
16	b_7	p_7

The decoded BAS value is valid if:

- the receiver is in frame and multiframe alignment, and
- the FAW in the same sub-multiframe was received with two or fewer bits in error.

Otherwise the decoded BAS value is ignored. When the receiver actually looses frame alignment, it may be advisable to undo any changes caused by the three previously decoded values as they may well have been erroneous even after correction.

3.2 Values of the BAS

The encoding of BAS is made according to a hierarchical attribute method. This consists of attribute class (8 classes), attribute family (8 families), attribute (8 attributes) and value (32 values). The first three bits of an attribute represent its number describing the general command or capability, and the other five bits identify the "value" - the specific command or capability.

The following attributes are defined in the Class (000) and Family (000):

Attribute	Significance
000	Audio Coding Command
001	Transfer Rate Command
010	Video and other Command
011	Data Command
100	Terminal Capability 1
101	Terminal Capability 2
110	Reserved
111	Escape Codes

The values of these attributes are listed and defined in Annex 1. They provide for the following facilities:

- transmission at various total rates and on single and multiple channels, on clear channels and on networks subject to restrictions to 56 kbit/s and its multiples;
- audio transmission, digitally encoded to various recommended algorithms;
- video transmission, digitally encoded to a recommended algorithm, with provision for future recommended improvement;
- low-speed data (LSD) within the I-channel, or TS1 of a higher rate initial channel;
- high-speed data (HSD) in the highest-numbered 64 kbit/s channel or time-slots (excluding the I-channel);
- data transmission within a multilayer protocol, either in the I-channel (MLP) or in capacity other than the I-channel (H-MLP);
- an encryption control signal;
- loopback towards the network for maintenance purposes;
- signalling for control and indications;
- a message system for, inter alia, conveying information concerning equipment manufacturer and type.

The command BAS attributes have the following significance: on receipt of a BAS command code in one (even) frame and its error-correcting code in the next (odd), the receiver prepares to accept the stated mode change beginning from the subsequent (even) frame; thus a mode change can be effected in 20 milliseconds. The command remains in force until countermanded (see Recommendation H.242, section 12). The bit positions occupied by combinations of BAS commands are exemplified in Figure 4(a to g).

The capability BAS attributes have the following significance: they indicate the ability of a terminal to receive and properly treat the various types of signal. It follows that having received a set of capability values from the remote terminal Y, terminal X must not transmit signals lying outside that declared range.

Values [0-7] of the attribute (111) are reserved for setting the class, and [8-15] for setting the family; the default value is (000) for both.

The next eight attribute values of the attribute (111) are temporary escape BAS codes of single byte extension (SBE). The last three bits of the temporary escape BAS form a pointer to one of eight possible escape BAS tables of 224 entries each (codes beginning with 111 are not used in the escape BAS tables). Then the next received BAS indicates the specific entry in the escape BAS table.

The value (111)[24] is the capability marker (see Recommendation H.242, section 2) which is followed by normal BAS codes, not by any escape values.

The last seven attribute values of the attribute (111) are of multiple byte extension (MBE) and are used to send messages as specified in the notes to the table in Annex 1.

3.3 Procedures for the use of BAS

The use of BAS codes is specified in Recommendation H.242.

Bit Number 7	Bit Number 8	Octet Number
1	FAS	1
2		2
:		:
8		8
9	BAS	9
:		:
16		16
17	18	17
19	20	18
:	:	:
143	144	80

FIGURE 4(a)/H.221

Bit numbering and position for 14.4 kbit/s LSD

Bit Number 1	2	3	4	5	6	7	8	Octet Number
1	2	3	4	5	6	7	FAS	1
:	:	:	:	:	:	:		2
:	:	:	:	:	:	:		:
50	51	52	53	54	55	56		8
57	58	59	60	61	62	63	BAS	9
:	:	:	:	:	:	:		:
:	:	:	:	:	:	:		:
106	107	108	109	110	111	112		16
113	114	115	116	117	118	119	SUB	17
120	121	122	123	124	125	126	CH.	18
:	:	:	:	:	:	:	8	:
:	:	:	:	:	:	:		:
554	555	556	557	558	559	560		80

FIGURE 4(b)/H.221

56 kbit/s LSD

Bit Number 1	2	3	4	5	6	7	8	Octet Number
1	2	3	4	5	6	7	FAS	1
:	:	:	:	:	:	:		2
:	:	:	:	:	:	:		:
50	51	52	53	54	55	56		8
57	58	59	60	61	62	63	BAS	9
:	:	:	:	:	:	:		:
:	:	:	:	:	:	:		:
106	107	108	109	110	111	112		16
113	114	115	116	117	118	119	120	17
121	122	123	124	125	126	127	128	18
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
617	618	619	620	621	622	623	624	80

FIGURE 4(c)/H.221

62.4 kbit/s LSD

Audio bit rate	Bit Number 1	2	3	4	5	6	7	8
G.711	MSB	...	...	...	...	...	...	LSB
G.722, 64 kb/s	H	H	L	L	L	L	L	L
G.722, 56 kb/s	H	H	L	L	L	L	L	
G.722, 48 kb/s	H	H	L	L	L	L		
16 kb/s	A1	A2						

A = audio bits
H = high-band audio
L = low-band audio

FIGURE 4(d)/H.221

Bit positions for audio

Initial channel								Additional channel							
Bit 1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
A1	A2	A3	A4	A5	A6	V1		V2	V3	V4	V5	V6	V7	V8	
A	..			..	A	V9		V10						V16	
:					:		FAS								FAS
							BAS								BAS
						V121		V122						V128	
						V129	V130	V131						V137	V138
						V139									V148
.					.	.									.
.					.	.									.
.					.	.									.
A	..			..	A	V759	..							..	V768

FIGURE 4(e)/H.221

Bit positions for video in 2B

TS1								TS2	TS3	TS4	TS5	TS6
A	A	A	A	A	A	A	F	V1 V8	V9 V16	V17 V24	D1 D8	D9 D16
							A	V25		V48	D17	D32
							S					
							B					
							A					
							S					
								V361		V384	D241	D256
							V	V386		V409	D257	
							V	V411				
							.	.				
							.	.				
							.	.		.	.	.
							.	.		.	.	.
							V	V1961. .		. . V1984	D1265. .	..D1280

FIGURE 4(f)/H.221

128 kbit/s HSD in H0 channel

Initial B-channel								2nd channel		3rd channel		4th channel		5th channel		6th channel
A	A	A	A	A	A	A	F	V1 V7	F	V8 V14	F	V15 V21	F	V22 V28	F	D1 D8
							A	V29	A		A	V42	A	V56	A	D9 D16
							S		S		S		S		S	
							B		B		B		B		B	
							A		A		A		A		A	
							S		S		S		S		S	
								V421						V448		D121 D128
							V	V450						V481		D129 D134
							V	V483						V514		D137 D144
							.	.						.		
							.	.						.		
							.	.						.		.
							.	.						.		.
							V	V2529. .						..V2560		D633...D640

FIGURE 4(g)/H.221

64 kbit/s HSD in 6 x 64 kbit/s channels

ANNEX 1

(to Recommendation H.221)

Definitions and tables of BAS values

The definitions of BAS values are given below, and the corresponding numerical values are listed in Tables A1/H.221 and A2/H.221.

Audio command values (000) - for bit position illustrations see Figure 4/H.221. Abbreviations "G.711" and "G.722" refer to CCITT Recommendations.

Neutral	Neutralized I-channel, containing only FAS and BAS; all other bits are to be ignored at the receiver
Au-off, U	No audio signal, no frame (Mode 10); all the I-channel is available for use under other commands[1]
Au-off, F	No audio signal, FAS and BAS in use (Mode 9); 62.4 kbit/s available for use under other commands
A-law, OU	G.711 audio at 64 kbit/s, A-law, no framing (Mode OU)[1]
A-law, OF	G.711 audio at 56 kbit/s, A-law, truncated to 7 bits in bits 1-7, with FAS and BAS in bit 8; bit 8 is set to zero at the PCM audio decoder (Mode OF)
μ-law, OU	G.711 audio at 64 kbit/s, μ-law, no framing (Mode OU)[1]
μ-law, OF	G.711 audio at 56 kbit/s, μ-law, truncated to 7 bits in bits 1-7, with FAS and BAS in bit 8; bit 8 is set to zero at the PCM audio decoder (Mode OF)
G.722, m1	G.722 7 kHz audio at 64 kbit/s, no framing (Mode 1)[1]
G.722, m2	G.722 7 kHz audio at 56 kbit/s, in bits 1-7 (Mode 2)
G.722, m3	G.722 7 kHz audio at 48 kbit/s, in bits 1-6 (Mode 3)
Au-40 k	Reserved for audio at less than 48 kbit/s (for example 40 kbit/s in bits 1-5)
Au-32 k	Reserved for audio at less than 48 kbit/s (for example 32 kbit/s in bits 1-4): the algorithm of "Au-16k" below may be extended to code a wider speech bandwidth at 32 kbit/s as a result of further studies
Au-24 k	Reserved for audio at less than 48 kbit/s (for example 24 kbit/s in bits 1-3)
Au-16 kbit/s	Audio at 16 kbit/s to Recommendation H.200/AV.254 in bits 1,2 (Mode 7)

[1] These attribute values designate unframed modes. In the receive direction reverting to a framed mode can only be achieved by recovering frame and multiframe alignment which might take up to two multiframes (320 ms).

Au-<16 k	Reserved for audio at less than 48 kbit/s (for example 8 kbit/s in bit 1)
Au-ISO-64/ 128/192/256	Audio to ISO standard at 64/128/192/256 kbit/s, in the lowest-numbered time-slots (other than TS1) of an H0 or greater channel
Au-ISO-384	Audio to ISO standard at 384 kbit/s in time-slots 2-7 of a channel greater than H0.

<u>Transfer-rate command values</u> (001)

<u>Note</u> - If the transfer-rate command is less than the available connected capacity, the information occupies the lowest-numbered channel(s)/time-slot(s).

64	Signal occupies one 64 kbit/s channel
2 x 64	Signal occupies two 64 kbit/s channels, with FAS and BAS in each .
3 to 6 x 64	Signal occupies three to six 64 kbit/s channels, with FAS and BAS in each
384	Signal occupies 384 kbit/s, with FAS and BAS in the first 64 kbit/s time-slot; the effective channel may be the whole of an H0 channel or the lowest numbered time-slots of an H11 or H12 channel
2 x 384	Signal occupies two channels of 384 kbit/s, with FAS and BAS in each
3 to 5 x 384	Signal occupies three to five 384 kbit/s channels, with FAS and BAS in each
1536	Signal occupies 1536 kbit/s, with FAS and BAS in the first 64 kbit/s time-slot. The effective channel occupies the whole of an H11 channel or the lowest numbered time-slots of an H12 channel
1920	Signal occupies 1920 kbit/s, with FAS and BAS in the first 64 kbit/s time-slot. The effective channel occupies the whole of an H12 channel
128/192/256	Signal occupies 128/192/256 kbit/s, with FAS and BAS in the first 64 kbit/s time-slot. The effective channel occupies the lowest numbered time-slots of an H0 or larger channel
512/768/ 1152/1472	Signal occupies 4512/768/1152/1472 kbit/s, with FAS and BAS in the first 64 kbit/s time-slot. The effective channel occupies the lowest numbered time-slots of an H11 or H12 channel
Loss-i.c.	Designated "Initial channel", especially used following loss of the channel previously so designated (see H.242, section 7.2.3)
channel #2-6	Numbering of additional channels - see section 2.7.1.

Video, encryption, loop and other commands (010)

video-off	No video; video switched off
H.261	Video on, to Recommendation H.261: video occupies all capacity not otherwise allocated by other commands; video cannot be inserted in the I-channel when var-LSD or var-MLP is in force; examples are given in Figure 4(e). Specifically, the video rate in initial B-channel (framed) or TSI is: 62.4 kbit/s - audio rate - (800 bit/s if ECS is ON) - (MLP rate if ON) - (LSD rate if ON)
vid-imp.(R)	Reserved for video on, to improved recommended algorithm
video-ISO	Video on, to ISO standard: video occupies the same capacity as stipulated above for the case of H.261 video
AV-ISO	Composite audio/video to ISO standard: the composite signal occupies the same capacity as stipulated above for the case of H.261 video
Freeze-pic.	Freeze-picture request (see Recommendation H.230, VCF)
Fast-update	Fast-update request (see Recommendation H.230, VCU)
encryp-on	ECS Channel active Note: when encryption is active, it applies to all information bits in all channels of the connection, except bits 1-24 of the SC in the I-channel and the FAS and BAS positions of the other channels; use of encryption in conjunction with MLP is for further study
encryp-off	ECS channel off (loopback requests are intended for use by maintenance staff)
Au-loop	Audio loop request (see Recommendation H.230, LCA)
Vid-loop	Video loop request (see Recommendation H.230, LCV)
Dig-loop	Digital loop request (see Recommendation H.230, LCD)
Loop-off	Loop off request (see Recommendation H.230, LCO)
6B-HO-comp	To provide for compatibility between terminals connected to single HO channel and six B-channel accesses, the least significant bits of the first 16 octets of all time-slots of the HO channel, except TS1, are not used; the HO terminal must discard these bits from the incoming signal on receipt of this code, and must set the same bits to "1" in the outgoing signal
Not-6B-HO	Negates the command "6B-HO-comp" Note: used, for example, in testing
restrict	To provide for operation on a restricted network, and for interconnection between a terminal on restricted and unrestricted networks: on receipt of this code, a terminal must treat the SC as being in bit 7 of the I-channel, and discard bit 8 of every other channel and/or time-slot; in the outgoing direction these bits are set to "1"

derestrict — On receipt of this code, a terminal must revert to "unrestricted network" operation, treating the SC as being in bit 8 of the I-channel

<u>LSD/MLP commands</u> (011) - for bit position illustrations see Figure 4/H.221.

#	these LSD rates are not allowed if ECS channel is in use
*	in restricted cases, the starred bit numbers are reduced by one
LSD off	LSD switched off
300	Low-speed data at 300 bit/s in SC, octets 38-40
1200	Low-speed data at 1200 bit/s in SC, octets 29-40
4800	Low-speed data at 4800 bit/s in SC, octets 33-80
6400	Low-speed data at 6400 bit/s in SC, octets 17-80#
8000	Low-speed data at 8000 bit/s in bit 7*
9600	Low-speed data at 9600 bit/s in bit 7* and octets 25-40 of SC
14400	Low-speed data at 14400 bit/s in bit 7* and octets 17-80 of SC#
16k	Low-speed data at 16 kbit/s in bit 6* and bit 7*
24k	Low-speed data at 24 kbit/s in bits 5*, 6* and 7*
32k	Low-speed data at 32 kbit/s in bits 4*-7*
40k	Low-speed data at 40 kbit/s in bits 3*-7*
48k	Low-speed data at 48 kbit/s in bits 2*-7*
56k	Low-speed data at 56 kbit/s in bits 1-7 (no framing in restricted case)
62.4k	Low-speed data at 62.4 kbit/s in bits 1-7 and octets 17-80 of SC. If ECS channel is in use the data rate is reduced to 61.6 kbit/s, but returns to 62.4 kbit/s if ECS channel is closed
64k	Low-speed data at 64 kbit/s in bits 1-8, no framing
var-LSD	Low-speed data occupying all I-channel capacity not allocated under other fixed-rate commands; cannot be invoked when other LSD is on, or when variable-MLP is on (may also be impractical when video is on in I-channel alone) Exact var-LSD rate: 62.4 kbit/s - audio rate - (800 bit/s if ECS in ON) - (fixed-MLP if ON)
dti(R)	three codes reserved for communicating the status of the data terminal equipment interfaces
MLP-off	MLP off in all channels

MLP-4k	MLP on at 4 kbit/s in octets 41-80 of SC
MLP-6.4k	MLP on at 6.4 kbit/s in octets 17-80 of SC; if ECS channel is in use, the data rate is reduced to 5.6 kbit/s in octets 25-80, but returns to 6.4 kbit/s if ECS channel is closed
var-MLP	MLP occupying all I-channel capacity not allocated under other fixed-rate commands: cannot be invoked when other MLP is on, or when variable-LSD is on (may also be impractical when video is on in I-channel alone); Exact var-MLP rate: 62.4 kbit/s - audio rate - (800 bit/s if ECS is ON) - (fixed-LSD if ON)

Audio capabilities (100)

neutral	neutral capability: no change in the current capabilities of the terminal
A-law	capable of decoding audio to Recommendation G.711, A-law
μ-law	capable of decoding audio to Recommendation G.711, μ-law
G.725-T1	Terminal Type 1 defined in Recommendation G.725, section 2
G.725-T2	Terminal Type 2 defined in Recommendation G.725, section 2
Au-16 kbit/s	capable of decoding audio, both to Recommendation H.200/AV.254 and Recommendation G.711
Au-ISO	capable of decoding audio to ISO standard at all rates up to 384 kbit/s

Video, MBE and encryption capabilities (101)

QCIF	Can decode video to QCIF picture format, but not CIF (see Recommendation H.261) - this code must be followed by one of the four MPI values below
CIF	Can decode video to CIF and QCIF formats (see Recommendation H.261) - this code must be followed by two MPI values, the first applicable to QCIF and the other to CIF format. Minimum picture interval (MPI) codes are as follows:
1/29.97	Can decode video, having a minimum picture interval of 1/29.97 seconds, to Recommendation H.261
2/29.97	Can decode video, having a minimum picture interval of 2/29.97 seconds, to Recommendation H.261
3/29.97	Can decode video, having a minimum picture interval of 3/29.97 seconds, to Recommendation H.261
4/29.97	Can decode video, having a minimum picture interval of 4/29.97 seconds, to Recommendation H.261
vid-imp(R)	Reserved for future improved recommended video algorithm

video-ISO	Can decode video to ISO standard
AV-ISO	Can decode composite audio/video signal to ISO standard
MBE-cap	Can handle multiple-byte extensions messages in the BAS position, those beginning with codes in the range (111)[25-31], in addition to other values
Esc-CF(R)	Reserved for capability to accept non-zero class/family escape codes
encryp.	Capable of handling signals on the ECS channel

<u>Transfer-rate capabilities</u> (100)

64, 384	Can accept signals only on one 64 kbit/s channel, one 384 kbit/s channel
2 x 64	Can accept signals on one or two 64 kbit/s channels, and synchronize them
...	...
6 x 64	Can accept signals on one to six 64 kbit/s channels, and synchronize them
2 x 384	Can accept signals on one or two 384 kbit/s channels, and synchronize them
...	...
5 x 384	Can accept signals on one to five 384 kbit/s channels, and synchronize them
1536/1920	Can accept signals on a 1536 kbit/s channel, a 1920 kbit/s channel
restrict	Can work only at p x 56 kbit/s, rate-adapted to p x 64 kbit/s by moving the SC to bit position 7 and setting bit 8 to "one" in every channel or time-slot; a constant "one", however, may be set in bit 8 if it is known by out-of-band signalling prior to the connection that the restriction exists; this code has the effect of forcing the remote terminal to work in the p x 56 kbit/s mode (see Annex 2)
6 B-H0-comp	Capable of acting upon the corresponding command
128/192/256	Capable of accepting the transfer rate specified by the corresponding command
512/768/ 1152/1472	Capable of accepting the transfer rate specified by the corresponding command

<u>LSD/MLP capabilities</u> (101)

300 (to 64k)	Can accept LSD at 300 bit/s (to 64 kbit/s) in the bit positions specified against the corresponding commands

var-LSD	Can accept LSD variable rate in the bit positions specified against the corresponding command
MLP-4k	Can accept MLP at 4 kbit/s in the SC
MLP-6.4k	Can accept MLP at up to 6.4 kbit/s in the SC
var-MLP	Can accept MLP at up to 64 kbit/s in the I-channel

Escape table values (111)

HSD	High-speed data: a 32-code table containing HSD capabilities and commands
H.230	Control and indications: a 32-code table with definitions in Recommendation H.230
start-MBE	First byte of (N+2) octet BAS message; the message format is: start-MBE//value of N (max=255)//N bytes
NS-cap	First byte of non-CCITT capabilities message; the message format is: NS-cap//value of N (max=255)//country code*//manufacturer code*//(N-4) bytes
NS-comm	First byte of non-CCITT command message; the message format is: NS-comm//value of N (max=255)//country code*//manufacturer code*//(N-4) bytes
cap-mark	Capability marker - the first item in a capability set - see Recommendation H.242, section 2
Data-apps	Applications within LSD/HSD channels: a 32-code table - see Table A3/H.221.

Note 1 - The value of N is coded by its binary representation.

Note 2 - The most significant bit of each MBE message byte is transmitted as the "b_0" bit of BAS.

HSD/H-MLP capabilities (111)10000-(101)

64k to 1536k	Can accept HSD at the specified rate in the bit positions specified against the corresponding commands
HSD-other	Reserved for other HSD rates
var-HSD	Can accept HSD variable rate in the bit positions specified against the corresponding command

* Country code consists of two bytes, the first being according to Recommendation T.35; the second byte and the terminal manufacturer code of two bytes are assigned nationally

H-MLP-62.4k Can accept MLP at 62.4 kbit/s in the bit positions specified against the corresponding command

H-MLP-r Can accept MLP at r=64/128/192/256/320/384 kbit/s in the bit positions specified against the corresponding command

var-H-MLP Reserved for capability to accept H-MLP variable rate in the bit positions specified against the corresponding command

<u>HSD/H-MLP commands</u> (111)10000-(011)

<u>Note</u> - In the cases of multiple channels, the term "highest-numbered time-slot" refers to the highest-numbered channel.

HSD-off HSD switched off; FAS and BAS restored in additional channels

64k HSD on, in highest numbered channel/time-slot; FAS and BAS are removed in the case of multiple B-channels

128/192/256k HSD on in highest-numbered time-slots of an H0 or greater channel

320k HSD on in highest-numbered time-slots of an H0 or greater channel

384k HSD on in highest-numbered H0 channel, or highest-numbered time-slots of a greater channel; FAS and BAS are removed in the case of multiple-H0 channels

HSD-other Reserved for other HSD rates

var-HSD Reserved for high-speed data occupying all capacity, other than in the I-channel, not allocated under other commands: cannot be invoked when other HSD is on, or when var-H-MLP is on (may also be impractical when video is on, the latter then being confined to the I-channel)

H-MLP-off H-MLP switched off (this does not affect I-channel MLP)

H-MLP-62.4 H-MLP on at 62.4 kbit/s, occupying second 64 kbit/s channel except FAS and BAS positions

H-MLP-64k
H-MLP-128k
H-MLP-192k
H-MLP-256k
H-MLP-320k
H-MLP on at 64/128/192/256/320 kbit/s in the lowest-numbered time-slots (other than TS1) of an H0 or greater channel

H-MLP-384k H-MLP on at 384 kbit/s in time-slots 2-7 of a greater channel than H0

var-H-MLP Reserved for MLP occupying all capacity, other than in the I-channel, not allocated under other commands: cannot be invoked when other MLP is on, or when var-HSD is on

<u>Note</u> - When the "restrict" command is in force the least significant bit of all octets covered by the HSD and H-MLP commands is set to "1", so the effective data rate is less than that indicated by the command.

<u>Applications within LSD/HSD channels - capabilities</u> (111)10010-(101)

ISO-SP baseline on LSD	Can accept ISO-still picture baseline mode on specified LSO rate
ISO-SP baseline on HSD	Can accept ISO-still picture baseline mode on specified HSD rate
ISO-SP spatial	Can accept ISO-still picture baseline - and spatial mode
ISO-SP progressive	Can accept ISO-still picture baseline - and progressive mode
ISO-SP arithmetic	Can accept ISO-still picture baseline - and arithmetic mode
Graphics cursor	Can handle graphics cursor data
Group 3 Fax	Can accept group 3 Fax
Group 4 Fax	Can accept group 4 Fax
V.120 LSD	Can accept V.120 terminal adaptation within an LSD channels
V.120 HSD	Can accept V.120 terminal adaptation within an HSD channel

<u>Applications within LSD/HSD channels - commands</u> (111)10010-(011)

ISO-SP on in LSD	ISO-still picture switched on in specified LSD
ISO-SP on in HSD	ISO-still picture switched on in specified HSD
Cursor data on in LSD	Cursor data switched on in specified LSD
Fax on in LSD	Fax switched on in specified LSD
Fax on in HSD	Fax switched on in specified HSD
V.120 LSD	V.120 switched on in specified LSD
V.120 HSD	V.120 switched on in specified HSD

TABLE A1/H.221

BAS numerical values

The column header gives the attribute designation as bits (b_0, b_1, b_2); the left-hand column gives the decimal value of bits [b_3, b_4, b_5, b_6, b_7]; for example, "channel #6" has the value (001)[10110]. All unassigned values are reserved, as are values marked (R).

	(000) audio command	(001) transfer rate command	(010) other command	(011) LSD/MLP command	(100) audio/ transfer-rate capability	(101) data/ video capability	(111) escape
[0]	neutral	64	video off	LSD off	neutral	var-LSD	
[1]		2x64	H.261	300	A-law	300	
[2]		3x64	vid-imp(R)	1200	μ-law	1200	
[3]		4x64	video-ISO	4800	G.725-T1	4800	
[4]	A-law, OU	5x64	AV-ISO	6400	G.725-T2	6400	
[5]	μ-law, OU	6x64		8000	Au-16kbit/s	8000	
[6]	G.722, m1	384	encryp-on	9600	Au-ISO	9600	
[7]	AU-off, U	2x384	encryp-off	14400		14400	
[8]	Note 1	3x384		16k	128	16k	
[9]	Note 1	4x384		24k	192	24k	
[10]		5x384		32k	256	32k	
[11]		1536		40k		40k	
[12]		1920		48k	512	48k	
[13]	Au-ISO-64	128		56k	768	56k	
[14]	Au-ISO-128	192		62.4k		62.4k	
[15]	Au-ISO-192	256		64k	1152	64k	
[16]	Au-ISO-256		freeze-pic	MLP-off	1B	MLP-4k	HSD
[17]	Au-ISO-384	loss i.c.	fast-update	MLP-4k	2B	MLP-6.4k	H.230
[18]	A-law,OF	chan.#2	Au-loop	MLP-6.4k	3B	var-MLP	Data-apps
[19]	μ-law,OF	chan.#3	Vid-loop	var-MLP	4B		(R-SBE)
[20]		chan.#4	Dig-loop		5B	QCIF	(R-SBE)
[21]		chan.#5	Loop-off	dti-1(R)	6B	CIF	(R-SBE)
[22]		chan.#6		dti-2(R)	restrict	1/29.97	(R-SBE)
[23]		512		dti-3(R)	6B-H0-comp	2/29.97	(R-SBE)
[24]	G.722,m2(Note2)	768			H0	3/29.97	cap-mark
[25]	G.722,m3(Note2)		6B-H0-comp		2H0	4/29.97	start-MBE
[26]	(Au-40k)	1152	No-comp 6B-H0		3H0	V-imp(R)	
[27]	(Au-32k)		restrict		4H0	video-ISO	
[28]	(Au-24k)		derestrict		5H0	AV-ISO	
[29]	Au-16kbit/s	1472			1472	esc-CF(R)	
[30]	(Au-<16k)				H11	encryp.	ns-cap
[31]	Au-off, F			var-LSD	H12	MBE-cap	ns-comm

Note 1 - These codes are listed in Recommendation G.725 with reference to an "application channel"; such a channel has not been defined, the concept having been superseded by that of LSD/MLP; therefore these codes should not be used.

Note 2 - These codes are listed in Recommendation G.725 with reference to "data"; however, the nature of such data (video, LSD, MLP, ECS) must be specified by further commands (001), (010), (011).

TABLE A2/H.221

HSD/H-MLP numerical values

Escape table reached by BAS (111)[16]

The column header gives the attribute designation as bits (b_0,b_1,b_2); the left-hand column gives the decimal value of bits [b_3,b_4,b_5,b_6,b_7]. All unassigned values are reserved, as are values marked (R).

	capabilities (101)	commands (011)
[0]		HSD-*off*
[1]	var-HSD(R)	var-HSD(R)
[2]	H-MLP-62.4	H-MLP-62.4
[3]	H-MLP-64	H-MLP-64
[4]	H-MLP-128	H-MLP-128
[5]	H-MLP-192	H-MLP-192
[6]	H-MLP-256	H-MLP-256
[7]	H-MLP-320	H-MLP-320
[8]	H-MLP-384	H-MLP-384
[9]		
[10]		
[11]		
[12]		
[13]	var-H-MLP(R)	var-H-MLP(R)
[14]		H-MLP-off
[15]		
[16]		
[17]	64k	64k
[18]	128k	128k
[19]	192k	192k
[20]	256k	256k
[21]	320k	320k
[22]	384k	384k
[23]	512k(R)	512k(R)
[24]	768k(R)	768k(R)
[25]	1152k(R)	1152k(R)
[26]	1536k(R)	1536k(R)
[27]		
[28]		
[29]		
[30]		
[31]		

TABLE A3/H.221

Numerical values for applications in LSD/HSD channels

Escape table reached by BAS (111)[18]

The column header gives the attribute designation as bits (b_0,b_1,b_2); the left-hand column gives the decimal value of bits [b_3,b_4,b_5,b_6,b_7]. All assigned values are reserved, as are values marked (R).

	capabilities (101)	commands (011)
[0]	ISO-SP baseline on LSD	ISO-SP on in LSD
[1]	ISO-SP baseline on HSD	ISO-SP on in HSD
[2]	ISO-SP spatial	
[3]	ISO-SP progressive	
[4]	ISO-SP arithmetic	
[5]		
[6]		
[7]		
[8]		
[9]		
[10]	Graphics cursor	Cursor data on in LSD
[11]		
[12]		
[13]		
[14]		
[15]		
[16]	Group 3 Fax	Fax on in LSD
[17]	Group 4 Fax	Fax on in HSD
[18]		
[19]		
[20]	V.120 LSD	V.120 LSD
[21]	V.120 HSD	V.120 HSD
[22]		
[23]		
[24]		
[25]		
[26]		
[27]		
[28]		
[29]		
[30]		
[31]		

ANNEX 2

(to Recommendation H.221)

Frame structure for interworking between a 64 kbit/s terminal and a 56 kbit/s terminal

1. Sub-channel arrangement

Bit number

1	2	3	4	5	6	7(SC)	8		
							1	1	Octet Number
S	S	S	S	S	S	FAS	1	.	
u	u	u	u	u	u		1	8	
b	b	b	b	b	b		1	9	
-	-	-	-	-	-	BAS	1	.	
C	C	C	C	C	C		1	16	
h	h	h	h	h	h		1	17	
a	a	a	a	a	a	(ECS)	1	.	
n	n	n	n	n	n		1	24	
n	n	n	n	n	n		1	25	
e	e	e	e	e	e		1	.	
l	l	l	l	l	l		1	.	
#	#	#	#	#	#	#	1	.	
1	2	3	4	5	6	7	1	80	

Note - C1, C2, C3 and C4 in the FAS are computed for the 160 septets, or 1120 bits.

2. Operation of the 64 kbit/s terminal

The transmitter fills the eighth sub-channel with "1", while the receiver searches FAS at every sub-channel.

3. Restriction against some communication modes

Since the interworking bit rate becomes 56 kbit/s, the transmission modes using more than 56 kbit/s are forbidden (receivers ignore these command BAS codes). Facilities using the original seventh sub-channel move to the sixth sub-channel.

4. Audio Command Codes (000) - the following are applicable instead of those in Annex 1.

Neutral	Neutralised I-channel, containing only FAS and BAS; all other bits are to be ignored at the receiver
Au-off, U	No audio signal, no framing; bits 1-7 of the I-channel are available
Au-off, F	No audio signal, FAS and BAS in use; 54.4 kbit/s available for use under other commands
A-law, U7	G.711 audio at 56 kbit/s, A-law truncated to 7 bits, no framing (Mode 0U)

A-law, F6	G.711 audio at 48 kbit/s, A-law truncated to 6 bits, with FAS and BAS in bit 7
μ-law, U7	G.711 audio at 56 kbit/s, μ-law truncated to 7 bits, no framing (Mode 0U)
μ-law, F6	G.711 audio at 48 kbit/s, μ-law truncated to 6 bits, with FAS and BAS in bit 7
G.722, U8	not possible to transmit 8 bits per octet
G.722, U7	G.722 7kHz audio in bits 1-7, 56 kbit/s (unframed)
G.722, F6	G.722 7 kHz audio at 48 kbit/s, in bits 1-6 (Mode 3)
Au-16 kbit/s	Audio at 16 kbit/s to Recommendation H.200/AV.254 in bits 1,2 (Mode 7)
[other]	All other values reserved

The following (000) values are assigned maintaining the same number of audio bits per octet between the 64 kbit/s and 56 kbit/s environments:

[0] Neutral
[6] not possible
[7] Au-off, U
[18] A-law, U7

[19] μ-law, U7
[20] A-law, F6
[21] μ-law, F6
[24] G.722, U7

[25] G.722, F6
[29] Au-16 kbit/s
[31] Au-off, F

2. Recommendation H.230

FRAME-SYNCHRONOUS CONTROL AND INDICATION SIGNALS FOR AUDIOVISUAL SYSTEMS

CONTENTS

1. Introduction

Digital audiovisual services are provided by a transmission system in which the relevant signals are multiplexed onto a digital path. In addition to the audio, video, user data and telematic information, these signals include information for the proper functioning of the system. The additional information has been named "control and indication" (C&I) to reflect the fact that while some bits are genuinely for "control", causing a state change somewhere else in the system, others provide for indications to the users as to the functioning of the system.

The C&I may be categorized into three groups:

a) call control - these are treated in Recommendations of the Q-Series

b) transmission frame-synchronous, or otherwise requiring rapid response

c) conference, data, and Telematic control not requiring frame synchronism, governed by the multilayer protocol (MLP) of Recommendation H.200/AV.270.

This Recommendation concerns only those C&I coming in category b) which includes a simplified set of conference C&I for multipoint connections of simple terminals.

2. Procedures

There are two procedures: some frame-synchronous C&I are provided for directly as BAS codes in Recommendation H.221, while the remainder require the use of an escape code.

2.1 C&I codes provided in Recommendation H.221

The following codes, whose functions are defined in section 3, are provided in Recommendation H.221:

- VCF, VCU (procedures for use in multipoint calls according to Recommendation H.200/AV.243);
- LCV, LCD, LCA, LCO (for maintenance - no standardized procedures).

In each case the code is transmitted in the BAS position at an appropriate time.

2.2 Other C&I codes

All frame-synchronous C&I codes not listed in section 2.1 are transmitted by a sequence involving the BAS positions in two consecutive sub-multiframes. In the first, the code (111)10001 is transmitted. In the second, the code defined in Table 1/H.230 is transmitted.

It should be noted that only one symbol is transmitted by this method - the code in the subsequent sub-multiframe is again treated as a normal BAS code.

3. Definitions of C&I symbols

The full definitions of these symbols are set out below and code values in Table 1/H.230. [The first letter of the alphabetic code-name indicates the type; the second is C for command, I for indication; the third is for the specific function.]

3.1 C&I related to video

VIS Video Indicate Suppressed: this symbol is used to indicate that the content of the video channel does not represent a normal camera image. The video encoder may be without video input or an electronically-generated pattern may have been substituted.

VIA Video Indicate Active: complementary to VIS. The video source is the only one, or, in the case that more video sources are to be distinguished, it is that designated "video #1".

VIA2 Equivalent to VIA, but designating "video #2" as the source.

VIA3 Equivalent to VIA, but designating "video #3" as the source.

VIR Video Indicate Ready-to-Activate: this symbol is transmitted by a terminal whose user has decided not to send video unless he will also receive video from the other end.

VCF Video Command "Freeze-Picture Request": this symbol may be transmitted prior to the "video-off" mode switch, to prepare the video decoder for this event. This symbol is also transmitted by a multipoint control unit (MCU) prior to video switching. On receipt, a terminal video decoder should complete updating of the current video frame but subsequently display the frozen picture until receipt of the freeze-picture release control which is embedded in the video.

VCU Video Command "Fast Update Request": this symbol is transmitted by an MCU after performing a video switch. It may also be transmitted by a terminal at the start of communication when the video decoder is first ready to receive. On receipt, the terminal video encoder should enter the fast-update mode at its earliest opportunity.

3.2 C&I related to audio

AIM Audio Indicate Muted: this symbol is used to indicate that the content of the audio channel does not represent a normal audio signal. The audio encoder may be without audio input or an electronically-generated tone may have been substituted.

AIA Audio Indicate Active: complementary to AIM.

3.3 C&I for maintenance purposes

LCV Loopback Command, "Video Loop Request": on receipt of this symbol, a terminal must connect the output of the video decoder to the input of the video encoder.

LCD Loopback Command, "Digital Loop Request": on receipt of this symbol, the terminal must disconnect the output of the multiplexer from the outgoing path, replacing it with the input to the demultiplexer. In the case of multiple B or HO connections, loopback is activated in each connection.

LCA Loopback Command, "Audio Loop Request": on receipt of this symbol, the terminal should if possible connect the output of the audio decoder to the input of the audio encoder.

LCO Loopback Command Off: on receipt of this symbol, the terminal must disconnect all loops and restore audio and data paths to their normal condition.

3.4 C&I related to simple multipoint conferences not using MLP

Note - Some of the following codes may be cancelled by transmission of appropriate codes as listed in Table 1/H.230 but not separately defined here.

MCV Multipoint Command Visualization-Forcing: transmitted by a terminal to force an associated MCU to broadcast its video signal. (Used to transmit the picture of a chairman or VIP, alternatively to hold a picture source during the transmission of graphics.)

MIV Multipoint Indication Visualization: transmitted by an MCU to indicate to a terminal that its video signal is being seen by other terminals (otherwise known as 'On-air' indication).

MCC Multipoint Command Conference: transmitted by an MCU to a terminal. The terminal receiving MCC must make its outgoing transfer rate equal to its incoming transfer rate, and its outgoing audio rate equal to its incoming audio rate.
Note - The command could also be used to invoke an on-screen user indication.

MCS Multipoint Command Symmetrical Data-transmission: transmitted by an MCU when setting up data broadcasting. On receipt, a terminal must prepare itself for data reception and ensure, by mode change if necessary, that its outgoing data channel occupies the same capacity as its incoming data channel. A terminal in receipt of MCS cannot initiate data broadcasting.

MCN Multipoint Command Negating MCS: transmitted by an MCU at the completion of data broadcasting. On receipt, a terminal must close any outgoing data channel which it has opened as a result of the previous reception of MCS. Following the end of data reception and the receipt of MCN, a terminal is permitted to initiate data broadcasting.

MIL Multipoint Indication Loop: an MCU has had its ports externally looped. The topic is for further study.

MIZ Multipoint Indication Zero-communication: transmitted by an MCU to a terminal for information, with the meaning that no other terminals are yet connected to the MCU.

MIS Multipoint Indication Secondary-status: transmitted by an MCU to a terminal for information, with the meaning that since other terminals of higher capability are participating in the conference-call, this terminal will not necessarily receive all the signals that are sent to those other terminals - see Recommendation H.200/AV.243.

MCA Multipoint Command Assign-token: possession of the token gives the holding terminal the right to give the MCU certain commands - see Recommendation H.200/AV.243.

MCT Multipoint Command Token-claim: sent by a terminal to the MCU. The MCU accedes to this claim if the token is unassigned or has been released.

MCR Multipoint Command Release-token: sent to the MCU by the terminal holding the token to give the MCU the authority to reassign the token to another terminal when/if it receives MCT.

4. Requirements for C&I

The C&I functions are defined such that, under various appropriate circumstances, the audiovisual system will operate in a fault-free manner and also such that sympathetic presentation to users is possible. Some functions must therefore be mandatory, others optional. This section, together with the categorization in Table 1/H.230, clarifies the circumstances under which C&I functions are mandatory.

CM denotes "conditionally mandatory": if the terminal (or MCU) is capable of entering the given state, then it must transmit the given code and, when leaving that state, the complementary code. If it has no such capability it can ignore both.

M denotes "mandatory" for all equipments of either terminal or MCU type

X denotes "non-mandatory": on receipt of such a code, it may be unrecognized, or recognized but not acted upon, or recognized and acted upon, entirely at the discretion of the manufacturer or user.

NA denotes that the code is not applicable in that case.

It will be noted that there are only a few mandatory requirements on most terminals. All audiovisual terminals must recognize and obey the command to make or break the digital loopback, and video loopback if they have video capability. All terminals having a video capability must also obey fast-update, freeze-picture, and MCS/MCN, otherwise there will be system misoperation on a multipoint call.

TABLE 1/H.230

Code first 3 bits	last 5 bits in decimal form	Value	TRANSMIT Term.	TRANSMIT MCU	RECEIVE Term.	RECEIVE MCU	Reference for procedures
(000)	[0,1]	Reserved					
	[2]	AIM	CM	CM	X	X	Section 3.2
	[3]	AIA	CM	CM	X	X	
	[4-15]	Reserved					
	[16]	VIS	CM	CM	X	X	Section 3.1
	[17]	VIA	CM	CM	X	X	Section 3.1
	[18]	VIA2	X	NA	X	X	H.320, AV.312
	[19]	VIA3	X	NA	X	X	H.320, AV.312
	[20-30]	Reserved					
	[31]	VIR	X	NA	X	NA	H.320
(001)	[0]	MCC	NA	M	M	NA	H.200/AV.243
	[1]	cancel-MCC	NA	M	M	NA	" "
	[2]	MIZ	NA	M	X	NA	" "
	[3]	cancel-MIZ	NA	M	X	NA	" "
	[4]	MIS	NA	M	X	NA	" "
	[5]	cancel-MIS	NA	M	X	NA	" "
	[6,7]	Reserved					
	[8]	MCT	X	NA	NA	M	" "
	[9]	MCR	X	NA	NA	M	" "
	[10]	MCA	X	NA	NA	M	" "
	[11-15]	Reserved					

[16]	MCV	X	NA	NA	M	H.200/AV.243
[17]	cancel-MCV	X	NA	NA	M	" "
[18]	MIV	NA	M	X	NA	" "
[19]	cancel-MIV	NA	M	X	NA	" "
[20]	MCS	NA	M	M	NA	" "
[21]	MCN	NA	M	M	NA	" "
[22-30]	Reserved					
[31]	MIL	NA	NA	NA	M	" "

(111) All codes forbidden

Code values listed in Recommendation H.221, Annex 1	VCF	X	M	M	NA	H.221
	VCU	X	M	M	NA	"
	LCV	NA	NA	CM	NA	"
	LCA	NA	NA	X	X	"
	LCD	NA	NA	M	X	"
	LCO	NA	NA	M	X	"

3. Recommendation H.242

SYSTEM FOR ESTABLISHING COMMUNICATION BETWEEN AUDIOVISUAL TERMINALS USING DIGITAL CHANNELS UP TO 2 MBIT/S

CONTENTS

1. Introduction

This Recommendation should be associated with Recommendations G.725 (System Aspects for the use of the 7 kHz Audio Codec within 64 kbit/s), H.221 (Frame Structure for 64-1920 kbit/s Channels in Audiovisual Teleservices) and H.230 (Frame-synchronous Control and Indication Signals for Audiovisual Systems).

A number of applications utilizing narrow (3 kHz) and wideband (7 kHz) speech together with video and/or data have been identified, including high quality telephony, audio and videoconferencing (with or without various kinds of Telematic aids), audiographic conferencing and so on. More applications will undoubtedly emerge in the future.

To provide these services, a scheme is recommended in which a channel accommodates speech, and optionally video and/or data at several rates, in a number of different modes. Signalling procedures are required to establish a compatible mode upon call set-up, to switch between modes during a call and to allow for call transfer.

Some services will require only a single channel, which could according to the procedures in this Recommendation be B (64 kbit/s), H0 (384 kbit/s), H11 (1536 kbit/s) or H12 (1920 kbit/s). Other services will require the establishment of two or more connections providing B or H0 channels: in such cases the first established is called hereafter the "initial channel" while the others are called "additional channels". Unless otherwise specified, all references to Frame Alignment Signal (FAS), Bit Rate Allocation Signal (BAS) and Service Channel (SC) refer to the initial channel or, in the case of a higher-order channel, to the time-slot No. 1 of this channel.

All audio and audiovisual terminals using G.722 audio coding and/or G.711 speech coding or other standardized audio codings at lower bit rates should be compatible to permit connection between any two terminals. This implied that a common mode of operation has to be established for the call. The initial mode might be the only one used during a call or, alternatively, switching to another mode can occur as needed depending on the capabilities of the terminals. Thus, for these terminals an in-channel procedure for dynamic mode switching is required.

The following paragraphs develop these considerations and describe recommended in-channel procedures.

2. Terminal capabilities

The procedures in this Recommendation are intended to ensure that only those signals are transmitted which can be received and appropriately treated by the remote terminal, without ambiguity. This requires that the capabilities of each terminal to receive and decode be known to the other terminal. Some capabilities are defined with a hierarchical structure: a terminal with capability value N is then also capable of all lower values. Where there is no hierarchy, then two or more codes of the same type may have to be transmitted in successive frames.

The following paragraphs define audio, video, transfer rate, and data rate capabilities of a terminal. It is not necessary that a terminal understand or store all incoming capabilities. Those which are not understood, or which cannot be used (because the terminal has no means to transmit corresponding information), can be ignored.

The total capability of a terminal to receive and decode various signals is made known to the other terminal by transmission (see section 5.1) of its capability set, consisting of the BAS-capability marker followed by all of the current capabilities. The codes are specified in Recommendation H.221, Annex 1; Table 1/H.242 (see section 12) summarizes the capabilities which may be included in a valid set. The transmission order is immaterial with the exception that video picture format values must be followed by minimum picture interval values.

Note - G.725 terminals send only a single capability value without a marker. The value is valid only if repeated at least once: this may be used to identify a G.725 terminal. Having so identified, the H.242 terminal should follow the procedures of G.725.

2.1 Audio capabilities

Audio capability values are defined in Recommendation H.221, Annex 1.

All audiovisual terminals intended for interregional operation should be capable of transmitting and receiving A- and μ-law G.711.

Normally, it is not necessary to transmit G.711 capabilities in a set containing other audio capabilities. Inclusion of just one value (A or μ) must be interpreted as a request not to send audio encoded to the other law - see section 6.3.1.

2.2 Video capabilities

Video capabilities are defined in Recommendation H.221, including:

- picture format: quarter-CIF, or both quarter-CIF and CIF;
- minimum picture interval (MPI): 1/29.97, 2/29.97, 3/29.97, 4/29.97 seconds.

The quarter-CIF value must be followed by one MPI value. The full-CIF value must be followed by two MPI values, the first applicable to quarter-CIF and the other to CIF.

2.3 Transfer rate capabilities

Transfer-rate capabilities are defined in Recommendation H.221.

The capability to receive a given number of multiple 64 kbit/s channel includes the capability to receive fewer 64 kbit/s channels. Similarly, the capability to receive a given number of H0 channels includes the capability to receive fewer H0 channels. In both cases the receiving terminal will synchronize the connected additional channels to the initial channel and maintain that synchronism throughout the period of connection.

All other ranges of capability must be signalled by inclusion in the capability set of more than one transfer rate capability code. For example, a terminal may list its transfer-rate capabilities as (2B and H0 and H11 and H12); in this case 1B capability is also implied.

2.4 Data capabilities

Data capabilities are defined in Recommendation H.221.

If a terminal is able to accept more than one data rate of whatever type (LSD, HSD, MLP, H-MLP) then all relevant values must be included in the capability set. Statement of one value does not include any other values.

2.5 Terminals on restricted networks: capability

A terminal connected to a network whose B-channels are effectively restricted to p x 56 kbit/s (p = 1 to 6), or whose channels at H0 or higher are restricted by ones-density considerations, must declare the capability value (100)[22] as given in Recommendation H.221. All terminals intended for interworking with terminals on restricted networks must have the capability to respond to this code according to Annex 2.

2.6 Encryption and extension-BAS capabilities

The capabilities are defined in Recommendation H.221.

3. Transmission

3.1 Transmission modes

Audio modes of operation are defined in Recommendation H.221, Annex 1, Audio Commands.

For analogue telephone terminals, it may be assumed that the speech signal is converted to PCM to G.711 at a digital network interface. These terminals are viewed as working in Mode 0U when connected to wideband speech terminals.

The video transmission is governed by the "video-on" and "video-off" commands. When switched on, the video signal occupies all of the capacity, both in the initial channel and in any additional channels, which is not specifically allocated to other signals by other commands. Thus different video bit rates will result from audio, transfer-rate, ECS and data commands the resultant video bit rate being: (transfer rate, less audio rate, less data rate if present, less encryption control channel if present, less FAS and BAS in all the channels/time-slots where they are present).

Transfer-Rate Modes are defined in Recommendation H.221, and specify the total capacity of the communication effective in the subsequent sub-multiframe.

Data modes are defined in Recommendation H.221, and specify only the bit rate and bit positions used for a user data signal. The protocol used for data applications is defined by the terminals, but see also section 9.

3.2 Establishment of compatible modes of operation

At the beginning of the communication phase of a call, all terminals start to work in Mode 0F. Terminals other than those limited to G.711 capability will then begin an initialization procedure.

This procedure (further described in section 6) consists of:

- the transmission of information concerning the capabilities of the respective terminals for receiving and decoding audio, video, transfer rate, data rates and other capabilities;

- the determination of a suitable transmission mode, consistent with the known capabilities of both terminals. An example is given in Appendix 4(a); in which the transmission mode is the same in both directions, but the H.242 procedures are equally applicable to systems in which asymmetric bidirectional communication is optimal (examples are surveillance - see Appendix 4(b) - and retrieval services);

- switching to this mode establishing additional channels if relevant.

The terminals connected to a call may change during the call. This may require re-initialization in order to identify the terminal type and to re-establish the desired mode of operation. In particular, this feature is used in mode 0 forcing, which is necessary in the case of a call transfer (see section 8).

4. Frame structure

The frame structure described in Recommendation H.221 is used for mode initialization and dynamic mode switching (see the following sections) and more generally to define the multiplex of the various bit streams (audio, video, data, encryption control signal, frame structure) within the frame.

Recommendation H.221 defines a Bit rate Allocation Signal (BAS) which is used *inter alia* to allocate sub-channels and to indicate the coding algorithm(s).

BAS codes are classified by the value of the first three bits which represent the BAS *attribute*: each attribute may therefore have up to 32 defined *values*.

Four BAS attributes are *commands*: they define the multiplex within the next and following sub-multiframes, as well as audio coding algorithm, and therefore command the distant receiver to treat the signals accordingly. The four attributes are independent; that is, a value of one attribute does not modify that of another.

Further BAS attributes are defined to signal terminal *capabilities* to the distant terminal. When received, these attributes do not directly affect the current transmission mode. However, they may lead to the initiation of a specific action to be carried out by the terminal. This feature is utilized in the mode initialization procedure and in the Mode 0 forcing procedure (see section 6).

The third bit of the H.221 Frame Alignment Signal (FAS) in odd frames of the initial channel, called the A-bit, is set to 1 on loss of frame or multiframe alignment, and is set to 0 on acquiring both frame and multiframe alignment (note). Consequently, a terminal which is receiving a framed signal with the A-bit set to 0 can assume that the distant terminal is able to act upon a change of BAS.

Note - A terminal having capabilities only for single-channel working and without encryption capability, does not need to seek and gain multiframe alignment since the latter serves for numbering and synchronizing multiple channels.

5. Basic sequences for in-channel procedures

Three signalling sequences are defined in this section. These sequences are used as the building blocks for the procedures defined in sections 6 and 7.

5.1 Capability exchange sequence A

The capability exchange sequence forces framing in both directions of transmission and the exchange of terminal capability codes. Either terminal may initiate the sequence and there is no problem caused by both doing so simultaneously or nearly simultaneously. Capability BAS should not be sent unnecessarily when the incoming signal is unframed.

The terminal X which initiates the capability exchange sequence must first switch to a framed mode if previously transmitting unframed; it then sets a timer T1 (value 10 seconds) and transmits its current capability set (see section 2) repetitively, or at least one complete set followed by the marker code (to indicate completion of the set); these capabilities will be one or more of the set listed in Table 1/H.242.

When Y first detects any incoming capability code except neutral (see section 5.3), it begins transmission of its own set of capability codes. This, of course, requires switching to a framed mode if transmission had been unframed. To ensure that each receives the complete set of capabilities of the other, they must continue repetitive transmission beyond the time they detect incoming A = 0 by at least one complete set and the marker code.

Note - See note on G.725 terminals in section 2.

There are three possible outcomes;

Outcome I: within the timer expiration period, multiframe alignment has been gained, the A bit is received with a value of 0 and the complete set of capability BAS codes of the distant terminal has been validated. In this case the sequence is completed successfully.

Note - If sequence A is initiated while incoming A = 0, repetition of the set is not necessary.

Outcome II: the timer has expired without multiframe alignment. In this case, the sequence failed.

Outcome III: the timer has expired with multiframe alignment achieved, but without either the validation of the A bit as 0 or the receiving of the complete set of the distant terminal's capability BAS codes (or both). In this case, the sequence is restarted. Outcome III should be notified to the user as a potential fault condition (which might, however, be in the remote terminal).

5.2 Mode switching sequence B

Mode switching is performed using BAS command codes, each being effective from the beginning of the even frame following the sub-multiframe in which the code is first transmitted. Mode switching is possible at any time during a communication, after the initialization procedure has been completed.

When the transmitting terminal signals the mode of operation this is valid from the next sub-multiframe. It is essential to note that transmitted signals must always be in accordance with the known capabilities of the remote terminal to receive and decode; in the absence of such knowledge only mode 0F or 0U (audio to Recommendation G.711) may be sent. If a change of capability, indicated in performing sequence A, has the result that the current mode is no longer receivable/decodable, there must be a switch as soon as possible to a mode which can be received and decoded.

The receiving terminal decodes and validates the BAS code, and switches its receive mode of operation accordingly. If for any reason a terminal receives a BAS command it cannot obey, a mode mismatch may result - see section 6.3.

In addition to switching of the audio mode, mode switching includes turning video off or on; the adoption/cessation of use of additional channels; the opening/closing of the encryption control channel; the opening/closing of a data channel.

The mode switching is in principle performed independently for the two transmission directions; some applications may be fundamentally asymmetric. For conversational services the terminal procedures will generally be such as to provide symmetrical transmission, though this is not mandatory.

5.3 Frame reinstatement sequence C (see Figure 1/H.242)

If terminal A is transmitting unframed but receiving framed, frame reinstatement consists in the insertion of FAS and BAS into the first 16 bits of the service channel, waiting for incoming A = 0; the overlaid frame can contain neutral BAS capability to avoid triggering a full capacity exchange.

A terminal A which is receiving unframed may wish the remote terminal B to reinstate framing: to do this, A must first itself reinstate framing if it is not already transmitting framed and then send the neutral BAS capability; B must respond by reinstating framing in order to return the neutral BAS capability and A = 0, and continuing this at least until it receives A = 0 itself.

FRAME REINSTATEMENT SEQUENCE "C"
(without consideration of restricted networks)

Is current outgoing signal
64 kbit/s Data? (Mode 10)

Send framed signal with suitable Data Command (62.4 kbit/s or less) - note that data is corrupted in the receiver until FAS is recovered at other end; interleave neutral-cap if relevant

Is current outgoing
signal 64 kbit/s Video?

Send framed signal with (000)[31] and (010)[1] - note that video is corrupted in the receiver until FAS is recovered at other end; interleave neutral-cap if relevant

Is current outgoing
signal Audio Mode 1?

Overlay framing without mode change; use (000)[6] and interleave neutral-cap if relevant

The current mode must
be PCM audio

Send Mode 0F with (000)[18 or 19] interleave neutral-cap if relevant

FRAME REINSTATEMENT SEQUENCE "C"
(application to restricted networks)

Is current outgoing
56 kbit/s Data?

Send framed signal with suitable Data Command (54.4 kbit/s or less) - note that data is corrupted in the receiver until FAS is recovered at other end; interleave neutral-cap if relevant

Is current outgoing
56 kbit/s Video?

Send framed signal with (000)[31] and (010)[1 or 2] - note that video is corrupted in the receiver until FAS is recovered at other end; interleave neutral-cap if relevant

Is current outgoing
signal 56 kbit/s G.722?

Overlay framing without mode change and interleave neutral-cap if relevant

The current mode must
be PCM audio

Send Mode 0F with (000)[18 or 19] and interleave neutral-cap if relevant

FIGURE 1/H.242

6. Mode initialization, dynamic mode switching and mode 0 forcing

Audiovisual terminals will be connected to digital networks where other kinds of terminals will also be connected: G.711 terminals but also data terminals, Telematic terminals, servers, etc. When compatibility between the different services involving those terminals is required an initialization procedure is necessary.

When automatic compatibility is required, a procedure based on the sequences defined in section 5 is used.

For call transfer or mode mismatch recovery it is necessary for terminals to operate in the common mode 0F and a mode 0 forcing procedure is required, again based on the sequences defined in section 5.

At the commencement of the call, after call transfer and after the procedure of section 6.3, there is a need for an initialization procedure to ensure that the two connected terminals can operate in the most suitable common mode.

6.1 Mode initialization procedure

6.1.1 Single channel

The initialization procedure begins as soon as a connection message is received from the network, or any indication meaning that the physical connection is established.

At the beginning of mode initialization, each terminal will start to transmit in mode 0F.

The receive part of the terminal should be in Frame Search and the receive audio is mode 0F. Sequence A is started.

Upon completion of sequence A according to outcome I, see Figure 2/H.242 (outcome Ia), sequence B will commence. The BAS code which is sent in sequence B is calculated from the knowledge of the capabilities of the local and distant terminals and is used to switch to a suitable working mode. This process may involve "terminal procedures" effecting choices made by the user or preset in the terminal. An example illustrating conformance to a defined teleservice is given in Recommendation H.320.

In the event of outcome II, the terminal will switch its transmission and reception to mode 0U. The receive part of the terminal should remain in Frame Search throughout the call.

In the event of outcome III, timer T1 is reset and the terminal remains within sequence A.

The initialization procedure is completed when both terminals have switched to the desired working mode(s).

6.1.2 Additional channels

A possibility of adding more channels is established from the capability exchange sequence. The calling terminal may then immediately begin establishing the additional connections. When each is established, it transmits only FAS and BAS on that channel, setting a timer T_a of value 10 seconds. Synchronization with the initial channel is performed according to Recommendation H.221, section 2.7. When the incoming A bits on additional channels are observed to be 0, mode switching to occupy sequentially numbered channels is initiated by an appropriate transfer-rate command BAS. If the timer T_a has expired without receiving A = 0, it is dealt with as a fault condition.

As the buffering process may involve the insertion of additional delay in the initial channel, which may already be carrying user information (speech, video, data), it may be necessary to make some provision for this interruption (e.g., short-term muting of audio output).

As additional channels achieve synchronization they are sequentially numbered using both FAS and BAS numbering as provided in Recommendation H.221.

An example of mode initialization on two channels is given in Appendix 1.

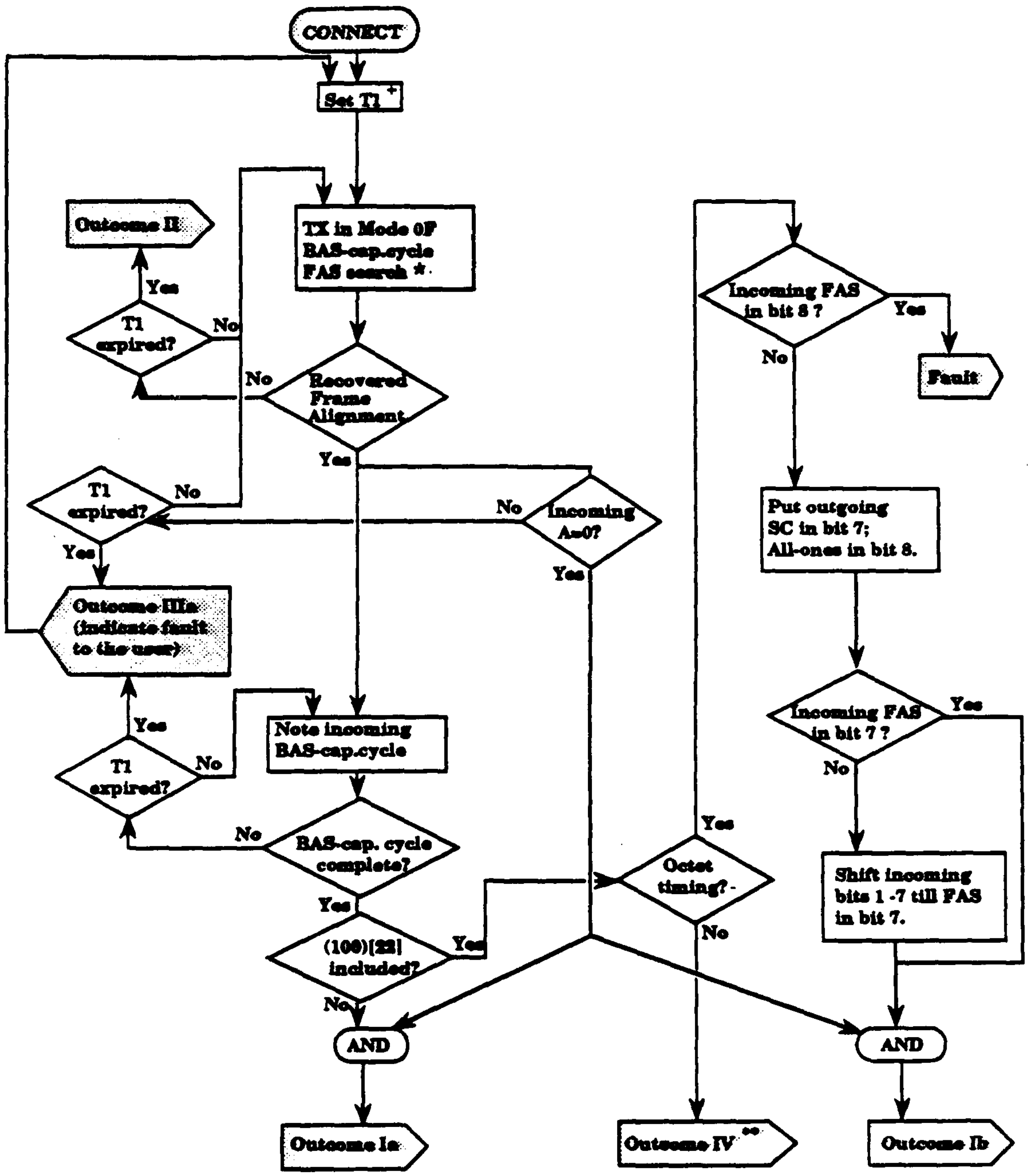
6.2 Dynamic mode switching (see Figure 3/H.242)

The mode switching procedure makes use of the frame structure specified in section 4 and of the sequences defined in section 5. It should be noted that all terminal receivers must remain in frame search throughout the call.

When the terminal is receiving in a framed mode, that is, it is capable of decoding bit A, mode switching should be delayed if the A bit is set to 1; eventually the Mode Mismatch Recovery procedure as described in section 6.4 might be used.

When the terminal X wishing to make a mode switch is receiving unframed signals, the capability exchange sequence may be used first to force the other terminal Y to a framed mode; hence terminal X can check for incoming A = 0. This use of sequence A is particularly necessary if X was previously transmitting unframed signals, since Y would not be in a position to deal with a mode change from X until it had regained frame alignment (see section 6.2.3). If X had previously been transmitting framed signals, the capability exchange sequence may be omitted on the assumption that if Y had unexpectedly lost frame alignment it would already have attempted a recovery procedure (see section 7).

INITIAL CAPABILITY EXCHANGE - GENERAL CASE



* Unless there is octet timing and it is certain that a restricted network is not involved, FAS should at this point be sought throughout the incoming signal.

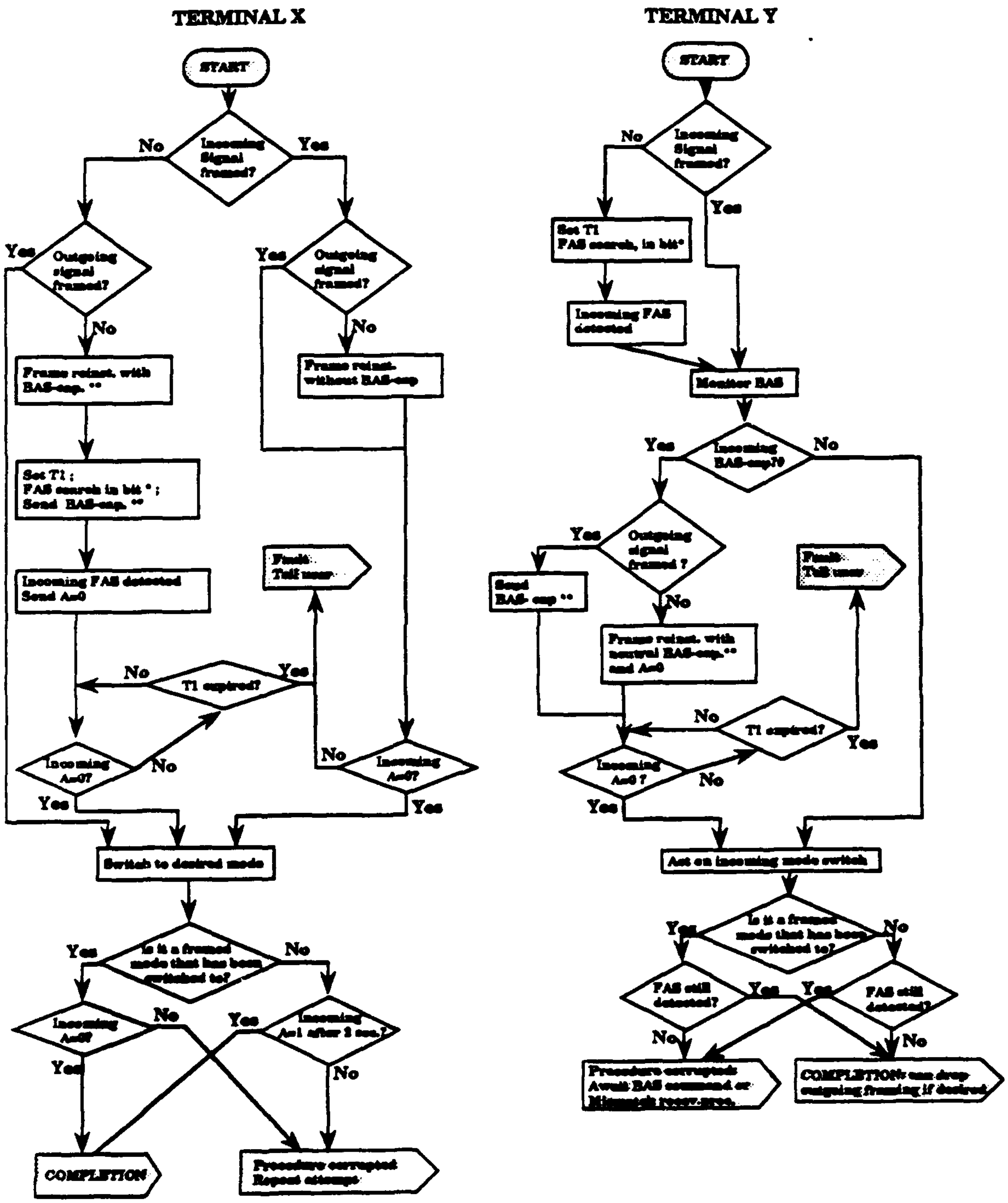
** Outcome IV: communication is impossible, because it is not known which bit is lost or stuffed; the terminal should so indicate to the user and wait.

\+ If the call is known to be inter-regional, it is advisable to mute the loudspeaker(s) until the audio decoder is set to the correct coding law.

FIGURE 2/H.242

MODE SWITCHING

Terminal X initiates the mode switch



* if byte-timed, search in appropriate bit
** neutral or complete BAS-cap cycle, depending on received BAS-cap.
# incoming BAS-cap causes outgoing frame reinstatement

T1506110-90

FIGURE 3/H.242

6.2.1 Dynamic mode switching from a framed mode to another framed mode

The basic sequence "Mode Switching" described in section 5.2 is used.

At the transmitting terminal, if a BAS command is transmitted to signal a new mode, the transmitter must operate in the appropriate mode from the first octet of the next sub-multiframe.

Similarly, at the receiving terminal, if the received BAS signals a new mode, the receiver must operate in the appropriate mode from the first octet of the next sub-multiframe.

6.2.2 Dynamic mode switching from a framed mode to an unframed mode

As above in 6.2.1, the basic sequence "Mode Switching" described in section 5.2 is used.

However, as the BAS for signalling an unframed mode is transmitted for a single sub-multiframe, a mode mismatch may occur in drastic error conditions. Optionally, a method may be used to improve the reliability of the switching: the new BAS value in the basic sequence "Mode Switching" is repeated three times; this will cause a temporary corruption of the least significant bit of the received information.

6.2.3 Dynamic mode switching from an unframed mode to another mode (framed or unframed)

The basic sequences "Frame Reinstatement" and "Mode Switching" are sequentially transmitted, the former including capability exchange if necessary.

6.3 Mode 0 forcing procedure - see Figure 4/H.242

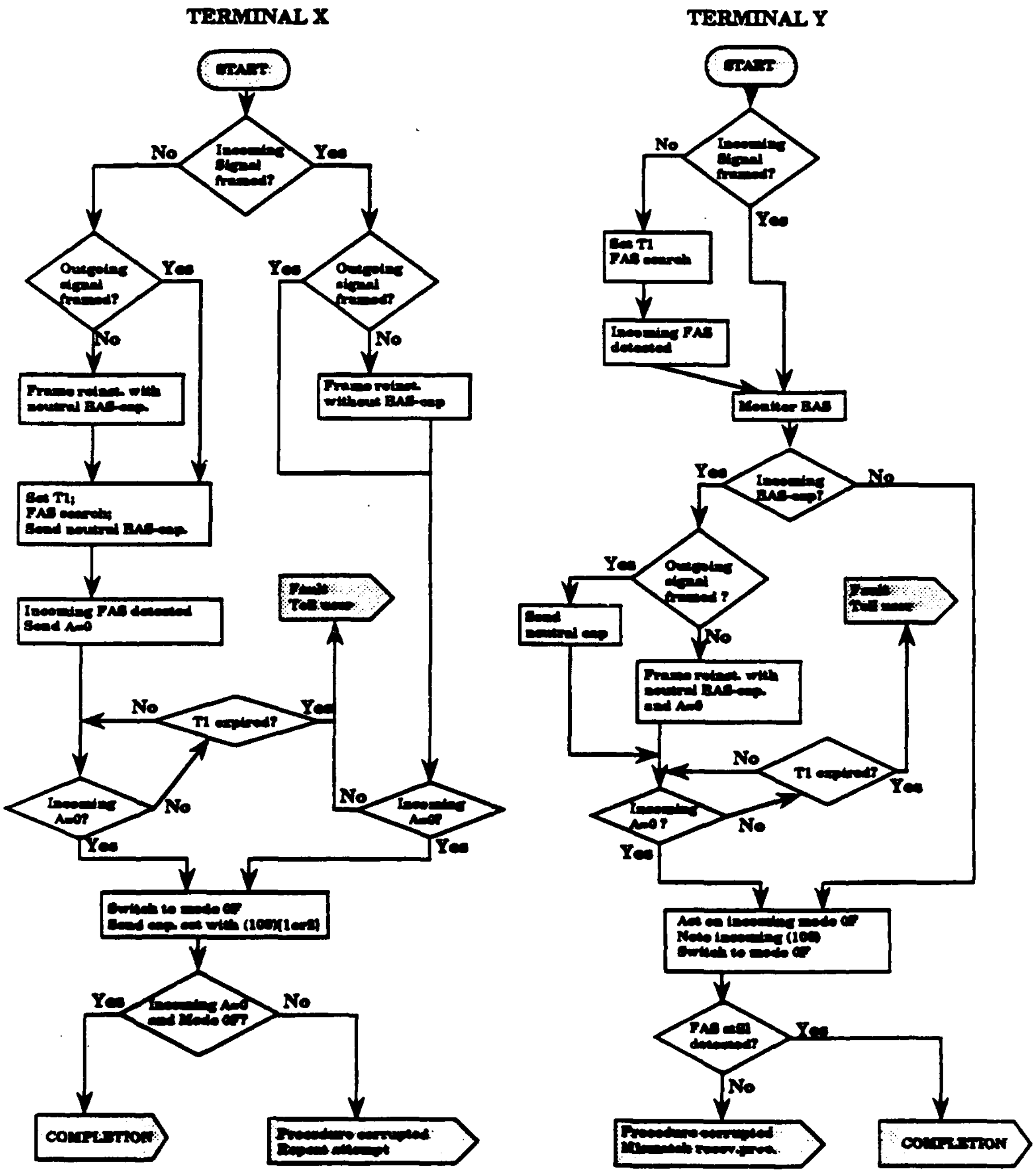
6.3.1 Single channel

Where it is necessary to ensure that both terminals are operating in Mode 0 (for instance before call transfer), this procedure is used.

The forcing terminal uses dynamic mode switching (section 6.2) with BAS audio command to switch to Mode 0F, followed by sequence A using BAS (100) indicating only G.711 audio capability. The value [1 or 2] appropriate to the terminal's own region is used in case the call is to be transferred to a local G.725 Type-0 terminal. On receipt of this, the remote terminal is obliged to switch to Mode 0F also using the indicated law for its encoder and decoder. The procedure is complete when the forcing terminal detects incoming Mode 0F. Changes of network configuration can now be implemented (see section 8).

MODE ZERO FORCING

Terminal X initiates the forcing



T1506120-90

FIGURE 4/H.242

6.3.2 Two or more channels

In this case the "Mode 0" forcing is applied to the initial channel only, and separate considerations apply to treatment of the additional channels. Three cases are considered here by way of guidance for the multiple-B case:

a) Additional channels dropped: this would be necessary, for example, prior to disconnection. The procedure is as for one channel, the forcing terminal declaring capability of PCM audio only with transfer rate capability of 1 x 64 kbit/s; this will result in mode switches successively to "data OFF", "video OFF" and audio mode OF or OU, such that all additional channels are vacated and can be disconnected.

b) Additional channels idle: this is the same as a) above, except that the forcing terminal makes no move to disconnect; the channels carry FAS, the multiframe number and the BAS indicating channel number; the content of the remainder of the idle channels is irrelevant.

c) Additional channels maintained active: this might be beneficial in some recovery procedures. The forcing terminal declares a capability of PCM audio plus transfer rate unchanged from its previous value, and then itself switches to the appropriate mode.

An example of Mode 0 forcing a) is given in Appendix 2.

6.4 Mode mismatch recovery procedure

In the case where mode mismatch has occurred, the Mode 0 forcing procedure may be used to establish a common working mode. Following this procedure, re-initialization can be achieved by using the mode initialization procedure.

7. Recovery from fault conditions

The provisions of this section are not wholly mandatory. In general it is expected that fault conditions will be rare and it may be uneconomical to provide elaborate recovery procedures to cover all eventualities. It is mandatory that proper indications of fault conditions be transmitted on the outgoing channel(s) - in particular, A must be set to 1 where appropriate conditions for A = 0 are not met. Other action to be taken on losing frame alignment, multiframe alignment, synchronism, or a connection, or on receiving incoming A = 1, is presented here for guidance.

7.1 Unexpected loss of synchronization or frame alignment

7.1.1 Loss of frame alignment in the initial channel

If a terminal unexpectedly loses frame alignment on its receive path, a timer T_3 is set (value for example 1 second) and incoming information is discarded if unintelligible. During this time the status of the framing in the receive direction is monitored:

a) If framing is recovered before the timer expires, the normal operation is resumed.

b) If framing is not recovered before the timer expires, the terminal goes to the Mode 0 forcing procedure followed by re-initialization.

7.1.2 <u>Loss of frame alignment of synchronization in an additional channel</u>

If a terminal unexpectedly loses synchronization (including that due to loss of frame alignment) on an additional channel, a timer T_3 is set, outgoing A-bit is set to 1 and incoming information discarded if unintelligible; if the loss of this information also causes information on other channels to become meaningless that also is discarded.

a) if synchronization is recovered before the timer expires, normal operation is resumed; this takes into account recoverable synchronization loss due to bit or synchronization errors on the transmission line;

b) if synchronization is not recovered before the timer expires, the mode 0 forcing procedure may be used.

7.2 <u>Recovery from loss of connection(s)</u>

Loss of a connection means that end-to-end transmission on that channel has been discontinued, so that all apparently received bits are meaningless. The receiver will, of course lose frame alignment and may follow the procedures of section 7.1. However, an indication may be available from the network (D-channel or otherwise) that the connection has been lost; in this case the procedures of this section are followed. It is assumed that connection loss is bidirectional; the case of loss in one direction only is for further study.

7.2.1 <u>Renumbering of channels</u>

This procedure is used for reconstructing the remaining normal additional channels when <u>one</u> additional channel breaks down.

i) Make the transmission mode of all channels into "framed".

ii) Vacate the sending additional channel(s).

iii) Renumber the additional channel(s).

iv) Wait for the synchronization establishment of the remote terminal and then expand communication onto the additional channels.

7.2.2 <u>Loss of an additional connection</u>

If any remaining channels are unframed (for example, data transmission) they must immediately have frame structure (according to Recommendation H.221) reimposed and maintained until conditions have returned to normal. The outgoing A-bit on additional channels is set to 1 if the incoming direction is unframed or out of sequence, or if synchronism has been lost.

If the lost channel was carrying part of a signal (such as encoded video) which also involved other channels, so that its loss renders the information in those other channels meaningless, then by dynamic mode switching those channels are vacated.

The next step is to renumber the available channels if appropriate, to obtain a continuous sequence; this is done using the procedure of section 7.2.1.

Dynamic mode switching is applied to re-establish the video or other transmission on the channels for which incoming A-bits are zero.

In the event that the lost channel be reconnected, it is added to the capacity in the same way as at the start of the call.

7.2.3 Loss of the initial connection

This results in the loss of the initial channel in both directions. Both terminals immediately regard #2 as the initial channel and transmit thereon the following BAS:

i) reinstatement of FAS and BAS in any unframed channels;

ii) transfer rate (001)[0 or 6] - code having the effect of vacating all additional channels; also audio command (000) unchanged from previous value;

iii) transfer rate (001)[17] on original second channel, indicating loss of original channel, and from next sub-multiframe original second channel substitutes for original initial channel; simultaneously any additional channels are renumbered in sequence;

iv) wait for confirmation that the synchronism at the remote terminal is retained/regained (all incoming $A_n = 0$);

v) expand communication onto all channels using appropriate transfer-rate command.

(Note - As a result of this procedure, sending and receiving initial channels may not be on the same connection.)

vi) the terminal tries to re-establish the lost channel.

8. Network consideration: call connection, disconnection and call transfer

8.1 Call connection

8.1.1 Initial channel

It is assumed that the terminals for switched network operation will have a signalling arrangement for originating calls over the network.

In the case that the network provides an indication that the connection is established (CONNECT-ACK message), the originating terminal will set its transmit and receive audio modes to PCM and begin the mode initialization procedure following the connection establishment indication. Where the network does not provide an indication of connection establishment the originating terminal will begin the mode initialization procedure immediately.

Upon answering a call the terminal will begin the mode initialization procedure.

Terminals for use on leased circuits may have a means for sending the alerting signal to the distant terminal and for answering the alerting signal. In this case, the sending of the alerting signal is equivalent to dialling and the foregoing procedures apply.

Whenever a terminal is manually reset, or recovers from a fault condition, the terminal will begin the Mode 0 forcing procedure of section 6.3. Then the terminal will begin mode initialization.

8.1.2 Additional channels

Call connection to provide additional channels may be initiated by one of the following:

a) manually;

b) on completion of the capability exchange sequence indicating mutual additional-channel capability;

c) at some time later than in b), prompted by user action.

The choice between these will depend on service provision and/or terminal procedures.

When the establishment of connection is known to the terminal, the mode initialization procedure of section 6.1.2 is applied.

During call establishment, an originating terminal should reserve additional channels by not answering incoming calls on those channels until it is determined whether the additional channels will be used in the connection. This prevents multiple call collisions and contention for the available channels. A network solution is under study.

8.2 Terminal disconnection

When a terminal disconnects from a call, the terminal must first initiate the Mode 0 forcing procedure, await completion of the procedure and then allow the actual disconnection of the call to occur.

8.3 Call transfer

As a consequence of the above, the terminal which continues to participate in a transferred call will be receiving in a PCM-forced state and therefore will be transmitting its capability set in framed PCM. When the transferred to terminal answers, mode initialization will occur in both directions.

8.4 Conferencing

Conferencing will be accomplished by means of a multipoint control unit (MCU). Each terminal will be connected to a port of the MCU by a switched connection or a leased circuit. Each connection between the terminal and the MCU is considered to be a point-to-point connection as far as call connection, terminal disconnection and call transfer procedures are concerned.

8.5 PCM format conversion

In the above procedures, no automatic method for establishing A-law or μ-law compatible PCM operation was defined.

At the beginning of the call, encoding and decoding by each terminal is to the law prevailing in its own region. The decoder must adapt to the coding law of the incoming signals. In a framed signal this will be clear from the BAS command; for unframed audio, signal analysis or local knowledge should be applied, and if this indicates that the other terminal is using a different coding law then the H.242 terminal should switch both its encoder and decoder to the coding law of the other terminal.

In the case where both terminals transmit framed signals, once the capability exchange is completed they may transmit in either PCM mode if desired.

Before call transfer, in the case where both terminals can transmit framed audio, the distant terminal's encoder and decoder must be forced by the relevant BAS capabilities and commands to the coding law of the region where the transfer is to take place.

9. Procedure for activaiton and de-activation of data channels

9.1 Data equipment not conforming to Recommendation H.200/AV.270

Each terminal must transmit a data-rate capability code (Recommendation H.221) for each data rate it is able to receive. This may be done during the capability exchange sequence at the start of the call or at a later time by initiating a new capability exchange.

A terminal may transmit data at any rate which has been indicated in the data-rate capability codes it has received from the other terminal. The appropriate data command (Recommendation H.221) is sent and in the following sub-multiframe the data transmission is commenced, occupying the bits within each frame defined in Recommendation H.221. However, at the time the data command is first sent, these bits must be unoccupied or contain only video information; therefore audio or any other signals must be removed from this part of the frame with the prior transmission of an appropriate command. In the case of occupancy by video information, commands are not available to reduce the video rate, but the video decoder continues to operate correctly on the lower flow of information. However, if the video rate is being made very low (for example, less than 30.4 kbit/s) or stopped altogether by the introduction of a data stream, it is advisable first to send "freeze-picture request", followed by the video "OFF" command.

The command "variable LSD" identifies as a data path the whole of the I-channel capacity not otherwise allocated by other commands; it must not be used when variable MLP is on, or when another LSD value is in force. If used while video is on, video is excluded from the I-channel.

At the conclusion of the data transmission the data "OFF" command is sent. If video is "ON" it will then occupy the freed bits in the next sub-multiframe and thereafter; otherwise those bits remain unoccupied until another command is sent.

At any time during data transmission the rate may be changed by an appropriate data command, subject to the provisions given above.

Note - In the case where 64 kbit/s HSD, for example, has been transmitted in the highest-numbered channel of a multiple-B channel connection, a slip during this data transmission would leave a misalignment when the HSD is turned off. To avoid corruption of video under these circumstances, it may be advisable to switch off the video stream before sending HSD-off, switching it on again as soon as A = 0 is received on the erstwhile data channel.

9.2 Equipment operating with an MLP according to Recommendation H.200/AV.270

Each terminal capable of operating with an MLP must transmit one of the MLP-capability codes. This may be done during the capability exchange sequence at the start of the call, or at a later time by initiating a new capability exchange.

When terminal X wishes to transmit MLP, it transmits MLP "ON" at the appropriate rate. Receiving the latter, terminal Y must establish an MLP channel at an appropriate rate (not necessarily the same rate) in the return direction.

The above provisions apply equally to the use of MLP on the I-channel, or in other channels or time-slots. Normally only one of these is required; however if both are in force, with appropriate commands, then a single MLP sub-channel at the combined rate may be interpreted - this would be specified within the appropriate service Recommendation (e.g. MLP rates of about 100 kbit/s on a 2B call).

To change the MLP rate, an appropriate MLP command is sent.

To discontinue use of the MLP, this matter may first be negotiated within the MLP itself; then one or both terminals transmit "MLP-OFF".

9.3 Simultaneous transmission of low-speed data and MLP

LSD and MLP may be active simultaneously, provided that no overlap is implied by the commands in force; however, variable LSD and variable MLP cannot coexist. No more than one LSD channel and one MLP channel may be active at any time (see also section 12).

10. Procedures for operation of terminals in restricted networks

Under study; the following paragraphs give preliminary considerations.

Terminals connected to a restricted network shall transmit the BAS capability "restricted" (100)[22] continuously when receiving an incoming A = 1 at the start of a call.

10.1 Network aspects

In this Recommendation the term "restricted network" applies to a network having restricted 64 kbit/s transfer capability, defined in Recommendation I.464 as "64 kbit/s octet-structured capability with the restriction that an all-zero octet is not permitted".

10.2 Reference connections

10.2.1 Case 1: 56 kbit/s, V.35 interfaces

Figure 5a shows a reference connection by a 56 kbit/s data service using V.35 interfaces. A 56 kbit/s clock is available at the V.35 interface; 8 kHz clock is not assumed. Figure 5c shows a reference connection, connected by 56 kbit/s network service with network clock.

10.2.2 Case 2: n x 56 kbit/s, V.35 interfaces

Figure 5b) shows a reference connection with more than two 56 kbit/s connections. Frame alignment will be according to H.221. Neither septet timing nor septet alignment is assumed. Figure 5d shows a multiple n x 56 kbit/s without septet alignment or septet timing.

10.2.3 Case 3: n x 64 kbit/s with octet timing and alignment

Figure 5e) shows a reference connection consisting of two visual telephones connected by facilities operating in a private line environment. Unrestricted mode of operation is not assumed.

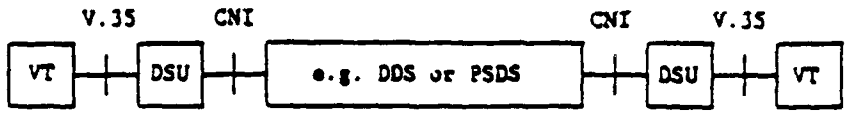


FIGURE 5a)

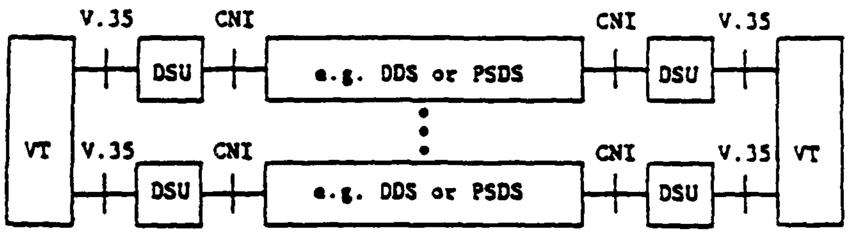


FIGURE 5b)

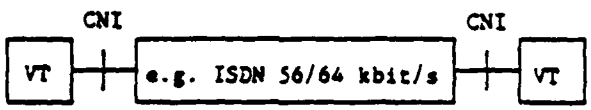


FIGURE 5c)

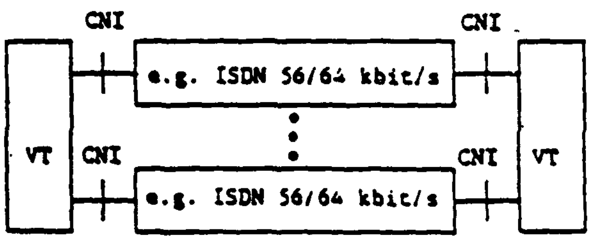


FIGURE 5d)

FIGURE 5e) T1502420-89

VT: Video Telephone
DSU: Data Service Unit
CNI: Customer Network Interface
DDS: Digital Data Service
PSDS: Public Switched Digital Service

10.2.4 Case 4: H0 (384 kbit/s) operation

When working in a restricted network a "1" shall be placed in the eighth bit position of every octet of every time-slot; the service channel is them in the seventh bit.

10.2.5 Case 5: 56 kbit/s satellite operation

For further study.

10.2.6 Case 6: 56 kbit/s interconnecting a 64 kbit/s network

A 64 kbit/s terminal will interwork with a 56 kbit/s terminal as a rate adapted data call over a 64 kbit/s bearer channel. The terminal connected to the 64 kbit/s connection will rate adapt according to H.221. In the case of a 64 kbit/s terminal connected to ISDN, the terminal may optionally be equipped to intercommunicate through an ISDN V.35 terminal adaptor. In any case, because the 56 kbit/s terminal cannot transmit correctly aligned septets, the terminal at the 64 kbit/s end cannot assume septet timing.

10.3 Transmission formats

10.3.1 Framing signal (56 kbit/s)

The transmission shall be arranged in 80 septet frames as specified in H.221.

10.3.2 Transmission formats (56 kbit/s operation)

In 56 kbit/s operation the septets of each 7 x 80 bit frame will be transmitted in order, most significant bit first at the 56 kbit/s rate. Septet alignment will be recovered from the frame alignment signal as specified in H.221.

10.3.3 n x 56 kbit/s operation

In n x 56 kbit/s operation each 56 kbit/s connection will be framed and transmitted separately. Septet timing will be recovered independently from the frame alignment signal of each channel, and the different delay between the channels will be compensated for on the basis of the multiframe numbering method specified in H.221.

The voice signal will be carried in the initial connection and video, graphics and auxiliary data may be carried in the initial and/or other connections.

10.3.4 n x H0 operation

In n x H0 operation each connection will be framed separately and differential delay between the channels will be compensated according to H.221.

10.3.5 Dynamic allocation within a primary-rate connection

Intelligent terminals may have a means for dynamically increasing or decreasing the bit rate during a connection. The means for controlling these allocations will be performed according to H.221. There may be a need to recover framing by extraction from the received signal independently.

10.4 Interworking between 56 kbit/s and 64 kbit/s terminals

In the worst case it must be assumed that neither terminal is aware (by means of a D-channel message or otherwise) that it is connected to a terminal of the other type; furthermore septet timing cannot be assumed at the 56 kbit/s end. At the 64 kbit/s end, byte timing is indispensable, since without this it cannot be known which bit (1 in every 8) will not be transmitted to the remote end (Figure 2/H.242, outcome IV).

Initially, terminal X (at 64 kbit/s) transmits FAS and capability-BAS on bit 8, on the false assumption that the remote terminal is also at 64 kbit/s. Frame search is carried out on the whole incoming signal; clearly, searching only on bit 8 will result in outcome II (Figure 2/H.242).

If frame alignment is found, and this may be in any bit position, given the lack of septet timing at the other end, then the fact of interworking with a 56 kbit/s terminal immediately becomes known from the capability BAS, which terminal Y must include in its capability BAS cycle. Terminal X immediately changes to transmitting FAS and BAS on bit 7, since bit 8 is the one which is not transmitted through the restricted networks. Initialization should then proceed as in section 6.1, with outcome Ib in Figure 2/H.242.

In the event that no frame alignment is found in any sub-channel, outcome II of section 6.1.1 applies.

Note 1 - All 56 kbit/s audiovisual terminals must transmit the appropriate capability BAS (100)[22] in every capability exchange.

Note 2 - Unless it is sure that they will never be required to interwork with 56 kbit/s networks, terminals manufactured for use on 64 kbit/s networks should preferably have the capability to search for frame alignment in all bit positions.

Note 3 - It may be advisable to mute audio output until incoming frame alignment has been achieved or a switch to unframed PCM has been decided upon.

10.5 Interworking between H0 or H11 terminals in restricted and unrestricted networks

At the start of the communication, the terminal on the restricted network transmits framed signals with the service channel in bit 7 of the I-channel and all "1"s in bit 8; the "restricted" capability BAS (100)[22] is sent. In the terminal on the unrestricted network, frame search is carried out on the whole incoming signal (or incoming TS1 if synchronization between H0/H11 framing and H.221 framing is maintained). When BAS (100)[22] is detected, a terminal immediately shifts the outgoing service channel to bit 7 and sets all "1"s on bit 8 of every time-slot.

All terminals intended for interworking with terminals connected to restricted networks must be capable of performing this procedure.

11. Procedure for use of BAS-extension codes

Recommendation H.221 provides for the attribute (111) for extension of the use of the BAS position in the subsequent sub-multiframe(s) for other purposes. There are 32 values of this attribute, the meanings of these being defined in H.221.

Note that the value (111)[24] is the capability marker (see section 2) which is followed by normal BAS codes, not by any escape values.

Value [0-15] are reserved for future extension of the scheme to include attribute class and family.

Values [16-23] are defined as single-byte extension (SBE); codes of SBE type may be transmitted at any time and to any terminal.

Value [18] gives access to a table of values specifying applications of a data channel (LSD or HSD). The application is active from the sub-multiframe following that in which the relevant specific application command BAS is transmitted. The closure of the data channel (using LSD/HSD-off) effectively closes the application.

All terminals must recognize the SBE attributes, at least to the extent of ignoring the subsequent code, whose meaning is not prescribed in this Recommendation. However, when (111)[17] is received, the subsequent code may be one of the mandatory values specified in Recommendation H.230. The ability of a terminal to use the content of other such codes is governed by other Recommendations. For example, Recommendation H.320 defines the requirements for visual telephone terminals to act upon some of the control and indication values.

Values [25-31] are of multiple byte extension (MBE); codes of MBE may only be transmitted to a terminal which has previously indicated its capability to receive MBE. It follows that a "non-CCITT capabilities" message may not be transmitted in the initial capability exchange, until the MPE-cap has been received. An example of the structure of MBE messages is given in Appendix 3.

12. Bit occupancy and the sequencing of BAS codes

In general, when there is no set procedure governing the sequence of BAS codes, priorities may be determined by the sending terminal. When there is no other demand for use of the BAS position, it is advisable to cycle through all the valid BAS commands, so that in the event of a temporary disturbance the proper mode will be restored as soon as possible thereafter.

Table 1/H.242 summarizes the BAS capabilities that can be simultaneously valid.

TABLE 1/H.242

Capability summary

Audio:	one or more values from A-law, μ-law, G.725-T1, G.725-T2, Au-16 kbit/s, Au-ISO
Video:	absent, or (QCIF plus one MPI value), or (QCIF + CIF plus two MPI values), and/or video-ISO and/or AV-ISO
Transfer rate:	absent (meaning rate = 64 kbit/s only* or up to four values: max. no. of 64, 384 kbit/s channels, 1536, 1920 kbit/s; and optionally any relevant values from {128, 192, 256, 512, 768, 1152, 1472 kbit/s}
Restricted network:	absent or present
Low-speed data (LSD):	absent or all relevant values
High-speed data (HSD):	absent or all relevant values
Low-speed MLP:	absent or all relevant values
High-speed MLP:	absent or all relevant values
Applications in data channel:	absent or all relevant values
Encryption:	absent or present
Multiple-byte extension:	absent or present

* When reducing the transfer-rate capability to 64 kbit/s from a higher value, the value "transfer-cap = 64 kbit/s" must be included.

The capability set consists of the capability market (111)[24] followed by all currently valid values, in any order; this may in turn be followed by a repetition of the set, or by the marker alone to indicate completion of the set prior to sending commands. No values should be repeated within a set. If it is desired to change the capability set during its transmission, the existing set must first be completed without change, followed by the marker alone and at least one BAS command before the new, changed set is started.

Table 2/H.242 summarizes the BAS commands that can be simultaneously valid.

TABLE 2/H.242

Command summary

Attribute	Alternative values (last value only is valid)	Default assumed	Comments
Audio (000)	[0, 4-7, 13-19, 24-31]	[18 or 19]	
Transfer rate (001)	[0-15, 23, 24, 26, 29]	[0]	
	[17]		see section 7.2.3
	[18-22]		additional channels only
Video and other (010)	[0-4]	[0]	
	[6, 7]	[7]	
	[16]		cancelled by command in video frame
	[17]		expires after fast update completed
	[18, 21]	[21]	
	[19, 21]	[21]	
	[20, 21]	[21]	
	[25, 26]	[26]	
	[27, 28]	[28]	
LSD and MLP (011)	[0-15, 31]	[0]	
	[16-19]	[16]	
HSD and H-MLP	[0, 17-22]	[0]	escape table (111)[16]
	[2-8, 13, 14]	[14]	

Only one value in each row can be in force at any one instant, up to 17 values on the initial channel (all the above values except (001)[18-22] apply only to the initial channel); however in practice many of the combinations are precluded by the fact that they would affect the same bits of the channel (for example, (011)[31] and (011)[19] cannot coexist).

A command remains in force until another from the same row is transmitted. A command must not be transmitted if to obey it would cause a simultaneous mode change on another row; in such a case the other row value must be changed first (for this purpose, a change of bit-rate of video or any of the variable data values does not constitute a mode change).

In general, unless specified otherwise, a BAS code which is invalid or which contravenes the provisions of this table, or otherwise indicates an impossible frame structure or system status, must not be transmitted.

The following notes serve to clarify the application of these rules to the multiplexing of audio, video and the various forms of data. Some examples relating to data transmission are given in Appendix 5.

a) Audio cannot penetrate into fixed rate Data (LSD or MLP) bit positions. It can expand its capacity into vacant or video or variable data bit positions. It can reduce its capacity within the audio bit positions currently occupied.

b) Video occupies all bit positions which are not assigned by other commands (ECS, Audio, LSD/MLP regardless of being fixed rate or variable rate).

Video can be turned on at any time even if the available capacity for video is zero at the corresponding sub-multiframe; (it may happen, for example, that video is switched on just before the variable rate LSD or MLP channel is closed); the decoder must not ignore "video on" even in this case, otherwise a mode mismatch occurs. However, if video capacity is less than about 30 kbit/s averaged over several sub-multiframes, it may not be practical.

It should be noted that video-off, (010)[0], is preferably preceded by freeze-picture request, (010)[16].

c) Fixed rate LSD/MLP cannot penetrate into Audio bit positions nor into fixed rate MLP/LSD bit positions. It can expand its capacity into vacant or video or variable MLP/LSD bit positions. It can reduce its capacity within the data bit positions currently occupied. As a combination, fixed rate LSD/MLP can occupy new bit positions which have previously been either vacant, video, variable rate MLP/LSD or occupied by the same type of fixed rate data.

d) Variable rate LSD/MLP occupies all bit positions which are not assigned by other fixed rate commands (ECS, Audio, fixed rate MLP/LSD). If video has been on, it is excluded when variable rate LSD or MLP is turned on. If variable rate LSD/MLP has been on, opening a variable rate MLP/LSD channel should be preceded by closing the existing variable rate LSD/MLP channel.

Variable rate LSD or MLP can be turned on at any time even if the available capacity for it is zero at the corresponding sub-multiframe; (it may happen, for example, that the variable MLP is switched on just before closing the LSD channel which has been occupying all the capacity other than audio); the decoder must not ignore "variable rate LSD or MLP on" even in this case, otherwise a mode mismatch occurs.

e) LSD/MLP rate may be changed without first closing the data channel - this applies equally to changes between fixed and variable rate. It is emphasized that there can only be one LSD and one MLP channel at any instant.

f) Capacity of video or variable LSD/MLP can be temporarily reduced to zero in a sub-multiframe as part of dynamic bit rate allocations. It is impractical, however, if that situation continues for a long time.

g) The rules for the use of HSD and H-MLP (in other than the I-channel) are identical to those given above for LSD and MLP in the I-channel.

13. Procedure for dealing with 6B-H0 interconnection

For further study.

14. Procedure for use of encryption control signal channel

Each terminal must transmit the encryption capability code if it is able to handle the ECS channel. No terminal may activate the channel without first receiving the corresponding capability code. Once an ECS capability code has been transmitted it cannot be cancelled by omission from a subsequent capability exchange. That is to say, a terminal having once received, stored and made use of an ECS capability code should assume continued validity until cancelled by the local user. Thus encryption can be discontinued by the users themselves but not by a third party tampering with the BAS-capability exchange.

The initiating terminal transmits the command "ECS Channel ON"; from the next multiframe it opens the 800 bit/s ECS channel defined in Recommendation H.221, whose use is specified in the Recommendation defining the encryption system (FAS, BAS and the ECS channel itself are in any case not encrypted).

When encryption has been turned off, the BAS command "ECS channel OFF" is used to close the ECS channel.

APPENDIX I

(to Recommendation H.242)

Initialization: Case of Videophone to Recommendation H.320, Type Xb_3

Underlined letters in the comments column correspond to points in the associated Figure A1.

SUCCESSIVE SUB-MULTIFRAMES AT TERMINAL "X" ONLY

TRANSMITTED					RECEIVED					
FAS, A-bit	BAS Attr.	BAS Value	Audio mode	Video rate	FAS, A-bit	BAS Attr.	BAS Value	Audio mode	Video rate	Comments
xx	xx	xx	xx	xx	xx	xx	xx	xx	xx	
F,1	(111)	[24]	0	(off)	xx	xx	xx	xx	xx	A cap-mark
F,1	(100)	[5]	0	(off)	xx	xx	xx	xx	xx	audio BAS-cap (1)
F,1	(100)	[4]	0	(off)	xx	xx	xx	xx	xx	audio BAS-cap (2)
F,1	(101)	[20]	0	(off)	xx	xx	xx	xx	xx	video Cap-QCIF
F,1	(101)	[24]	0	(off)	xx	xx	xx	xx	xx	MPI 3/29.97
F,1	(100)	[17]	0	(off)	xx	xx	xx	xx	xx	Transfer rate Cap 2B
F,1	(111)	[24]	0	(off)	xx	xx	xx	xx	xx	repeat cap-set
F,1	(100)	[5]	0	(off)	xx	xx	xx	xx	xx	
	(continue to cycle caps)				(searching for frame alignment)					about one transit?
F,1	(101)	[24]	0	(off)	xx	xx	xx	xx	xx	
F,1	(100)	[17]	0	(off)	F,1	(111)	[24]	0	(off)	B incoming cap-set
F,1	(111)	[24]	0	(off)	F,1	(100)	[5]	0	(off)	...
F,1	(100)	[5]	0	(off)	F,1	(100)	[4]	0	(off)	
F,1	(100)	[4]	0	(off)	F,1	(101)	[20]	0	(off)	...
F,1	(101)	[20]	0	(off)	F,1	(101)	[24]	0	(off)	...
F,1	(101)	[24]	0	(off)	F,1	(100)	[17]	0	(off)	...
F,1	(100)	[17]	0	(off)	F,1	(111)	[24]	0	(off)	cap-set complete
...	...	...	...	...	(searching for multiframe align.)					up to 320 ms
F,0	(101)	[24]	0	(off)	F,1	(100)	[17]	0	(off)	C mfa achieved, A=0
F,0	(100)	[17]	0	(off)	F,1	(111)	[24]	0	(off)	
...	...	...	...	...	(waiting for incoming A=0)					
F,0	(100)	[17]	0	(off)	F,1	(111)	[24]	0	(off)	
F,0	(111)	[24]	0	(off)	F,0	(100)	[5]	0	(off)	D incoming A=0
F,0	(100)	[5]	0	(off)	F,0	(100)	[4]	0	(off)	
F,0	(100)	[4]	0	(off)	F,0	(101)	[20]	0	(off)	...
F,0	(101)	[20]	0	(off)	F,0	(101)	[24]	0	(off)	...
F,0	(101)	[24]	0	(off)	F,0	(100)	[17]	0	(off)	...
F,0	(100)	[17]	0	(off)	F,0	(111)	[24]	0	(off)	
F,0	(111)	[24]	0	(off)	F,0	(100)	[5]	0	(off)	cap-set complete
F,0	(000)	[29]	0	(off)	F,0	(100)	[4]	0	(off)	E start mode switch
F,0	(010)	[1]	7	(off)	F,0	(101)	[20]	0	(off)	(Note 1)
F,0	(000)	[29]	7	46.4	F,0	(101)	[24]	0	(off)	
F,0	(010)	[1]	7	46.4	F,0	(100)	[17]	0	(off)	
F,0	(000)	[29]	7	46.4	F,0	(111)	[24]	0	(off)	
F,0	(010)	[1]	7	46.4	F,0	(100)	[5]	0	(off)	

..	...	...	...	...	(waiting for incoming mode changes)					
F,0	(010)	[1]	7	46.4	F,0	(101)	[24]	0	(off)	
F,0	(000)	[29]	7	46.4	F,0	(000)	[29]	0	(off)	F incoming switch
F,0	(010)	[1]	7	46.4	F,0	(010)	[1]	7	(off)	16 kbit/s audio
F,0	(000)	[29]	7	46.4	F,0	(000)	[29]	7	46.4	video ON
F,0	(010)	[1]	7	46.4	F,0	(010)	[1]	7	46.4	repeat valid
F,0	(000)	[29]	7	46.4	F,0	(000)	[29]	7	46.4	commands
(now deal with second B-channel, once connection is completed)										
FF,01	(010)	[1]	7	46.4	Fx,0x	(000)	[29]	7	46.4	G
FF,01	(000)	[29]	7	46.4	Fx,0x	(010)	[1]	7	46.4	
...	...	...	...	...	(searching for frame alignment on channel #2)					
FF,01	(010)	[1]	7	46.4	FF,01	(000)	[29]	7	46.4	H align. recovered
FF,01	(000)	[29]	7	46.4	FF,01	(010)	[1]	7	46.4	
...	...	...	...	...	(finding multiframe alignment and buffering to synchronize)					
FF,00	(010)	[1]	7	46.4	FF,01	(000)	[29]	7	46.4	I send A=0 on chan. #2
FF,00	(000)	[29]	7	46.4	FF,01	(010)	[1]	7	46.4	
...	...	...	...	...	(waiting for incoming A_2=0)					
FF,00	(010)	[1]	7	46.4	FF,00	(000)	[29]	7	46.4	J incoming A_2=0
FF,00	(001)	[1]	7	46.4	FF,00	(010)	[1]	7	46.4	start mode switch to
FF,00	(001)	[1]	7	108.8	FF,00	(000)	[29]	7	46.4	expand video (Note 1)
FF,00	(010)	[1]	7	108.8	FF,00	(010)	[1]	7	46.4	
FF,00	(000)	[29]	7	108.8	FF,00	(000)	[29]	7	46.4	
FF,00	(001)	[1]	7	108.8	FF,00	(010)	[1]	7	46.4	
(continue to cycle BAS commands)					(waiting for incoming mode changes)					
FF,00	(010)	[1]	7	108.8	FF.00	(001)	[1]	7	46.4	K incoming mode sw.
FF,00	(000)	[29]	7	108.8	FF.00	(001)	[1]	7	108.8	
(initialization completed)										

Note 1 - The modes selected for switching are governed by "terminal procedures" which in general depend on the application; in the present case of videophone service, the procedure is specified in Recommendation H.320.

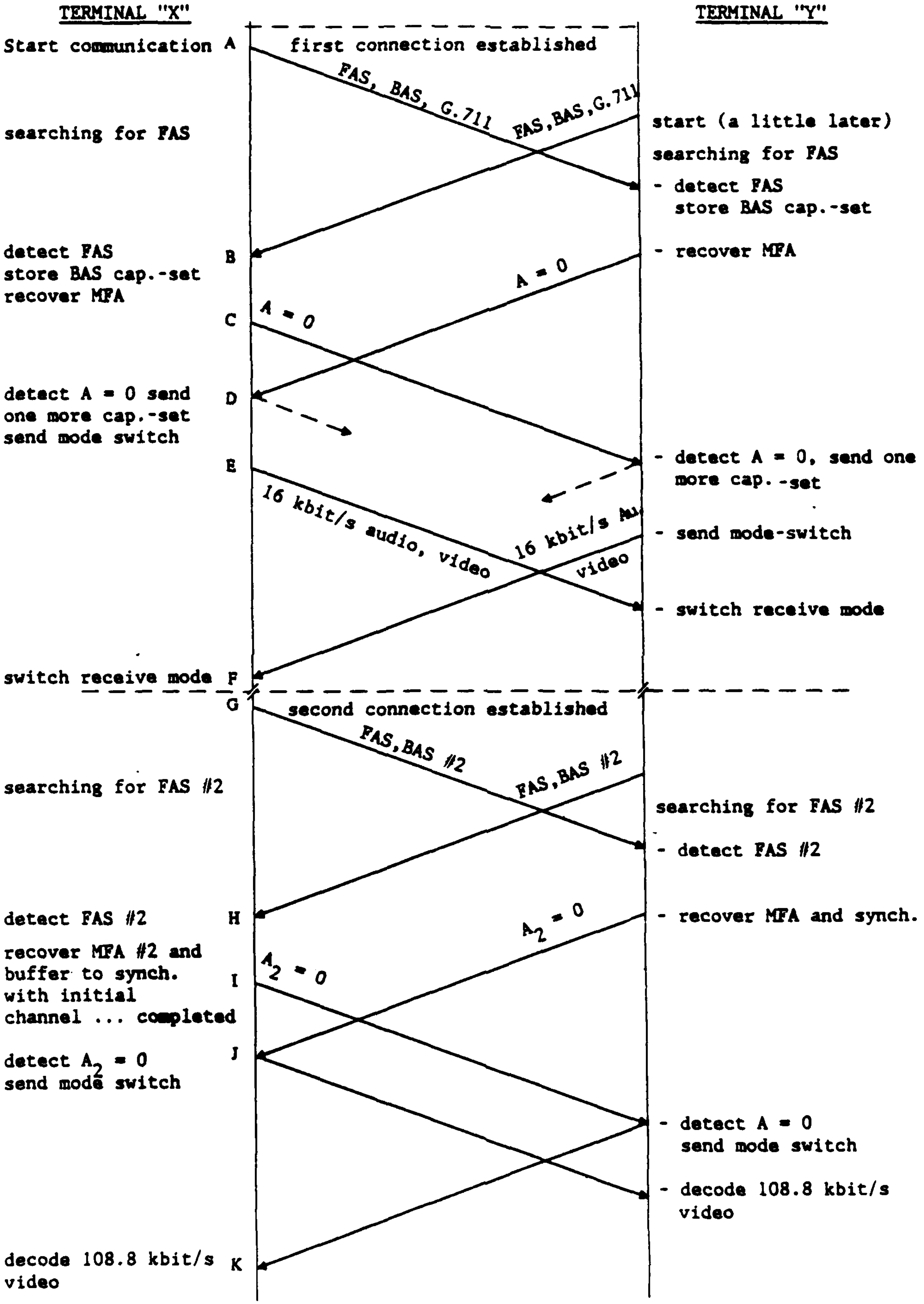


FIGURE A1/H.242

T1508070-90

APPENDIX II

(to Recommendation H.242)

Mode-0 forcing: Case of Videophone to Recommendation H.320, Type Xb_3

Underlined letters in the comments column correspond to points in the associated Figure A2.

SUCCESSIVE SUB-MULTIFRAMES AT TERMINAL "X" ONLY

TRANSMITTED					RECEIVED					
FAS, A-bit	BAS Attr.	Value	Audio mode	Video rate	FAS, A-bit	BAS Attr.	Value	Audio mode	Video rate	Comments
FF,00	(010)	[1]	7	107.6	FF,00	(000)	[29]	7	107.6	Video is ON (H.261)
FF,00	(000)	[29]	7	107.6	FF,00	(001)	[1]	7	107.6	Audio is 16 kbit/s
FF,00	(001)	[1]	7	107.6	FF,00	(011)	[2]	7	107.6	Transfer rate is 2 x 64
FF,00	(011)	[2]	7	107.6	FF,00	(010)	[1]	7	107.6	Data is ON at 1.2 kbit/s
FF,00	(010)	[1]	7	107.6	FF,00	(000)	[29]	7	107.6	
FF,00	(011)	[0]	7	107.6	FF,00	(001)	[1]	7	107.6	L Data to go off
FF,00	(010)	[0]	7	106.6	FF,00	(011)	[2]	7	107.6	Video to go off
FF,00	(001)	[0]	7	(off)	FF,00	(010)	[1]	7	107.6	Transfer rate 64 kbit/s
FF,00	(000)	[18]	7	off	FF,00	(000)	[29]	7	107.6	Audio A-law, OF
FF,00	(000)	[18]	OF	off	FF,00	(001)	[1]	7	107.6	
FF,00	(010)	[0]	OF	off	FF,00	(011)	[2]	7	107.6	
FF,00	(000)	[18]	OF	off	FF,00	(010)	[1]	7	107.6	
FF,00	(111)	[24]	OF	off	FF,00	(000)	[29]	7	107.6	M cap mark
FF,00	(100)	[16]	OF	off	FF,00	(001)	[1]	7	107.6	64 kbit/s-cap only
FF,00	(100)	[1]	OF	off	FF,00	(011)	[2]	7	107.6	A-law capability only
FF,00	(111)	[24]	OF	off	FF,00	(010)	[1]	7	107.6	cap-mark
(continue to cycle these caps)					(awaiting incoming mode change and cap. set)					
FF,00	(100)	[16]	OF	off	FF,00	(000)	[29]	7	107.6	
FF,00	(100)	[1]	OF	off	FF,00	(011)	[0]	7	107.6	N Incmg. data to go off
FF,00	(111)	[24]	OF	off	FF,00	(010)	[0]	7	106.6	Incoming video to go off
FF,00	(100)	[18]	OF	off	FF,00	(001)	[0]	7	(off)	Incoming chan. #2 off
FF,00	(100)	[1]	OF	off	FF,00	(000)	[18]	7	(off)	Incoming audio to be OF
FF,00	(010)	[0]	OF	off	FF,00	(111)	[24]	OF	off	
FF,00	(001)	[0]	OF	off	FF,00	(100)	[5]	OF	off	
FF,00	(000)	[18]	OF	off	FF.00	(100)	[4]	OF	off	
FF,00	(011)	[0]	OF	off	FF.00	(101)	[20]	OF	off	
FF,00	(010)	[0]	OF	off	FF.00	(101)	[24]	OF	off	
FF,00	(001)	[0]	OF	off	FF.00	(100)	[17]	OF	off	
FF,00	(000)	[18]	OF	off	FF.00	(111)	[24]	OF	off	

(continue to cycle all valid BAS commands)

The mode-0 forcing procedure is not complete: subsequent action depends on the "terminal procedure", according to the reason for performing the switch to mode zero.

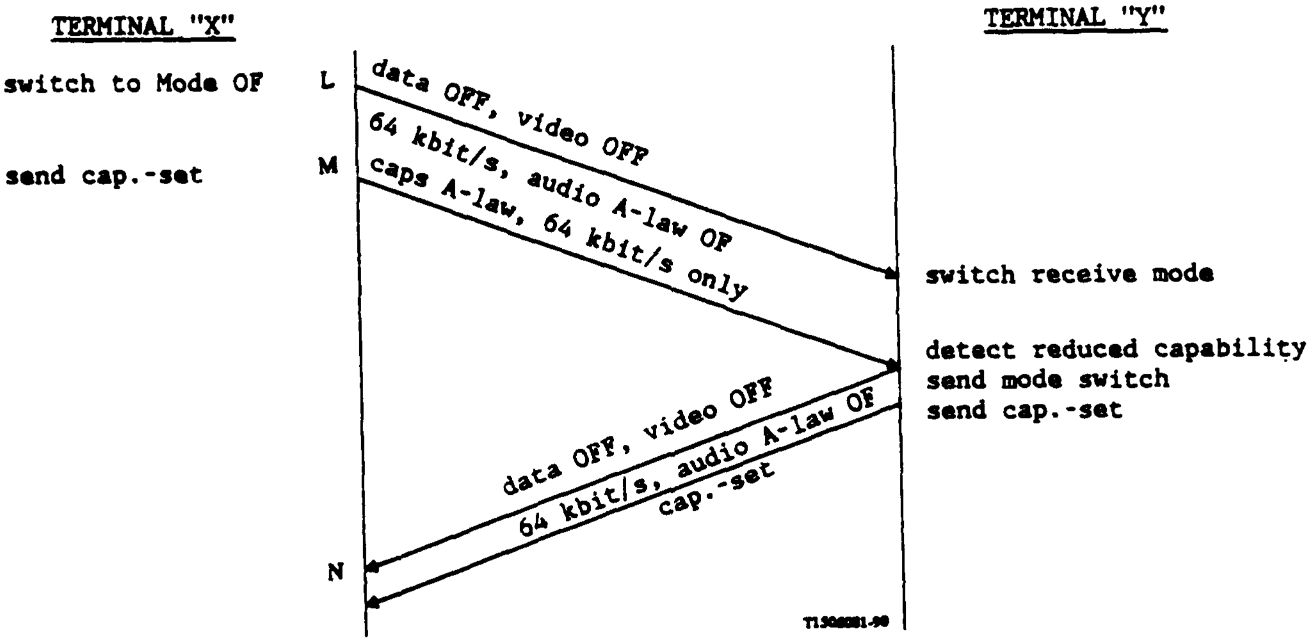


FIGURE A2/H.242

APPENDIX III

(to Recommendation H.242)

Example of use of message structure

SEND RECEIVE

1. Initial Capability Exchange, including MBE-cap

```
(111) [24]  Cap-mark
(100) [4]   Audio Type 2 (G.722, 56 kbit/s)
(100) [17]  2 x 64 kbit/s transfer rate
(101) [21]  CIF video capability
(101) [22]  1/29.97 MPI for QCIF
(101) [23]  2/29.97 MPI for CIF
(101) [31]  MBE-cap
(111) [16]  Set to escape table for HSD
(101) [17]  64 kbit/s HSD-cap
(111) [24]  Cap-mark, repetition of capability set
(100) [4]   Audio Type 2 (G.722, 56 kbit/s)
 ...   ...   ...
```

decode incoming BAS capabilities: these include (101)[31], so remote end can handle MBE codes

2. Subsequent Capability Exchange, including MBE capability message

```
(111) [24]  Cap-mark
(100) [4]   Audio Type 2 (G.722, 56 kbit/s)
(100) [17]  2 x 64 kbit/s transfer rate
(101) [21]  CIF video capability
(101) [22]  1/29.97 MPI for QCIF
(101) [23]  2/29.97 MPI for CIF
(101) [31]  MBE-cap
(111) [16]  Set to escape table for HSD
(101) [17]  64 kbit/s HSD-cap
(111) [30]  Start of non-CCITT capability message
(M)         Information will be M bytes
(byte 1)    Country code according to Recommendation T.35
(byte 2)    Country code
(bytes 3,4) Manufacturer code (Company XYZ)
(bytes 5-M) Type identity
(111) [24]  Cap-mark, repetition of capability set
(100) [4]   Audio Type 2 (G.722, 56 kbit/s)
 ...   ...   ...
```

incoming capability cycle now includes the same non-standard mode

3. Mode switch to non-standard mode using MBE command

```
(111) [30]  Start of non-CCITT command message
(N)         Information will be N bytes
(byte 1)    Country code according to Recommendation T.35
(byte 2)    Country Code
(bytes 3,4) Manufacturer code (Company XYZ)
(bytes 5-N) Type identity
```

The mode switch is effective from the sub-multiframe following that containing byte N.

APPENDIX IV

(to Recommendation H.242)

Examples of symmetrical and unsymmetrical transmission modes

a) Example of symmetrical transmission mode

	Audio	Video	Transfer rate	LSD	HSD	MLP
Capabilities of Terminal X	16 kbit/s	Yes	1B	1.2 kbit/s	-	No
Capabilities of Terminal Y	Type 2 +16 kbit/s	Yes	2B	1.2 kbit/s	-	Yes
Mode in X-to-Y direction	16 kbit/s	ON	1B	1.2 kbit/s	-	OFF
Mode in Y-to-X direction	16 kbit/s	ON	1B	1.2 kbit/s	-	OFF

b) Example of unsymmetrical transmission mode

	Audio	Video	Transfer rate	LSD	HSD	MLP
Capabilities of Terminal X	PCM	Yes	2B	1.2 kbit/s	No	No
Capabilities of Terminal Y	16 kbit/s	No	2B	56 kbit/s	No	No
Mode in X-to-Y direction	OFF	OFF	2B	56 kbit/s	-	OFF
Mode in Y-to-X direction	OFF	ON	2B	1.2 kbit/s	-	OFF

APPENDIX V

(to Recommendation H.242)

Examples relating to data transmissions

Transfer-rate 1B, audio at 48 kbit/s, no video or video off

MLP	LSD	Forbidden next commands (example)
4k	1200	#,LSD=4.8k/6.4k/14.4k and over, MLP=6.4k
4k	8k	Au=56k, #, LSD=4.8k/6.4k/14.4k and over
4k	var	#, LSD=4.8k/6.4k/14.4k and over, MLP=var
6.4*k	8k	Au=56k, #, LSD=300/1200/4.8k/6.4k/9.6k/14.4k and over
var	1200	#, LSD=16k and over/var, MLP=6.4k
var	6.4k	#, LSD=16k and over/var, MLP=4k/6.4k
var	9.6k	Au=56k, #, LSD=16k and over/var, MLP=6.4k

Transfer-rate 1B, audio at 16 kbit/s, no video or video off

MLP	LSD	Forbidden next commands (example)
4k	300	LSD=4.8k/6.4k/14.4k/48k and over, MLP=6.4k
4k	8k	Au=56k, LSD=4.8k/6.4k/14.4k/48k and over
4k	16k	Au=48k/56k, #, LSD=4.8k/6.4k/14.4k/48k and over
4k	var	#, LSD=4.8k/6.4k/14.4k/48k and over, MLP=var
6.4*k	8k	Au=56k, LSD=300/1200/4.8k/6.4k/9.6k/14.4k/48k and over
6.4*k	40k	Au=48k/56k, #, LSD=300/1200/4.8k/6.4k/9.6k/14.4k/48k and over
var	4.8k	#, LSD=48k and over/var, MLP=4k/6.4k
var	9.6k	Au=56k, #, LSD=48k and over/var, MLP=6.4k
var	16k	Au=48k/56k, #, LSD=48k and over/var

Transfer-rate 1B, audio at 16 kbit/s, video on

MLP	LSD	Forbidden next commands (example)
4k	1200	LSD=4.8k/6.4k/14.4k/48k and over, MLP=6.4k
4k	8k	Au=56k, LSD=4.8k/6.4k/14.4k/48k and over
6.4*k	8k	Au=56k, LSD=300/1200/4.8k/6.4k/9.6k/14.4k/48k and over

"video-on" may not be practical in these cases.

Transfer-rate 2B, audio at 48 kbit/s, video on

MLP	LSD	Forbidden next commands (example)
var	1200	LSD=16k and over/var, MLP=6.4k
var	4.8k	LSD=16k and over/var, MLP=4k/6.4k
var	9.6k	Au=56k, LSD=16k and over/var, MLP=6.4k
4k	8k	Au=56k, LSD=4.8k/6.4k/14.4k/16k and over

Transfer-rate 2B, audio at 16 kbit/s, video on

MLP	LSD	Forbidden next commands (example)
var	1200	LSD=48k and over/var, MLP=6.4k
var	4.8k	LSD=48k and over/var, MLP4k/6.4k
var	8k	Au=56k, LSD=48k and over/var
var	16k	Au=48k/56k, LSD=48k and over/var
4k	8k	Au=56k, LSD=4.8k/6.4k/14.4k/48k and over

* These rates are reduced by 800 bit/s when the ECS is active.

\# "video on" may not be practical in these cases.

4. Recommendation H.261

VIDEO CODEC FOR AUDIOVISUAL SERVICES AT p x 64 kbit/s

CONTENTS

The CCITT,

considering

(a) that there is significant customer demand for videophone, videoconference and other audiovisual services;

(b) that circuits to meet this demand can be provided by digital transmission using the B, H0 rates or their multiples up to the primary rate or H11/H12 rates;

(c) that ISDNs are likely to be available in some countries that provide a switched transmission service at the B, H0 or H11/H12 rate;

(d) that the existence of different digital hierarchies and different television standards in different parts of the world complicates the problems of specifying coding and transmission standards for international connections;

(e) that a number of audiovisual services are likely to appear using basic and primary rate ISDN accesses and that some means of intercommunication among these terminals should be possible;

(f) that the video codec provides an essential element of the infrastructure for audiovisual services which allows such intercommunication in the framework of Recommendation H.200;

(g) that Recommendation H.120 for videoconferencing using primary digital group transmission was the first in an evolving series of Recommendations,

appreciating

that advances have been made in research and development of video coding and bit rate reduction techniques which lead to the use of lower bit rates down to 64 kbit/s so that this may be considered as the second in the evolving series of Recommendations,

and noting

that it is the basic objective of the CCITT to recommend unique solutions for international connections,

recommends

that in addition to those codecs complying to Recommendation H.120, codecs having signal processing and transmission coding characteristics described below should be used for international audiovisual services.

Note 1 - Codecs of this type are also suitable for some television services where full broadcast quality is not required.

Note 2 - Equipment for transcoding from and to codecs according to Recommendation H.120 is under study.

1. Scope

This Recommendation describes the video coding and decoding methods for the moving picture component of audiovisual services at the rates of p x 64 kbit/s, where p is in the range 1 to 30.

2. Brief specification

An outline block diagram of the codec is given in Figure 1/H.261.

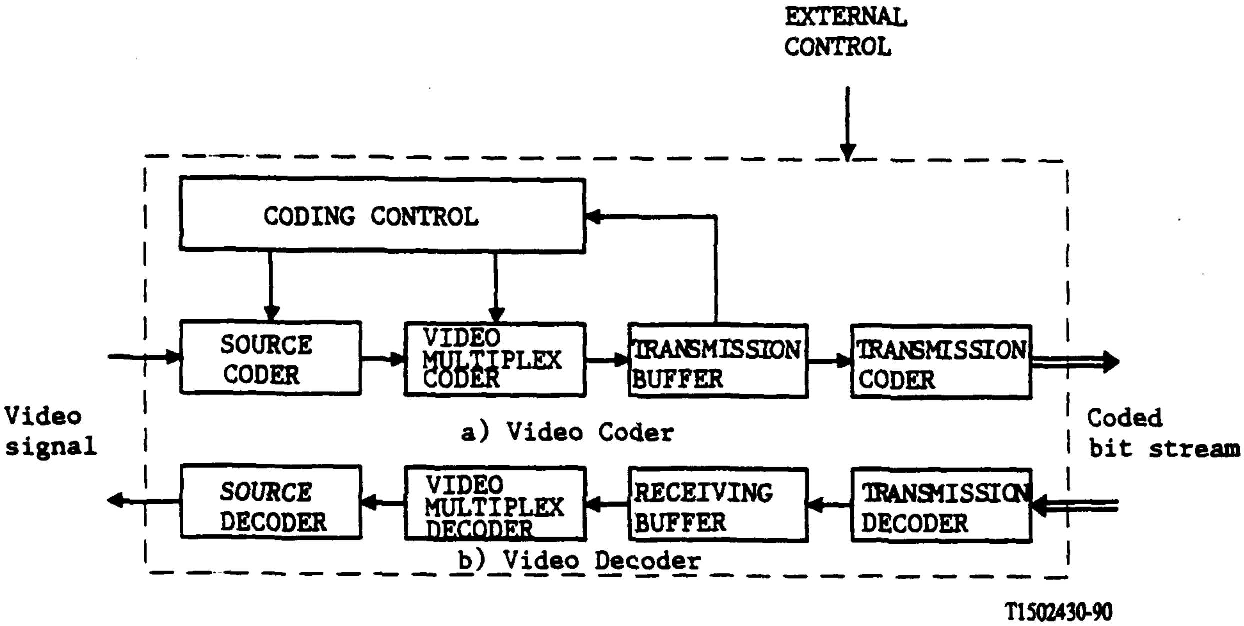


FIGURE 1/H.261

Outline block diagram of the video codec

2.1 Video input and output

To permit a single Recommendation to cover use in and between regions using 625- and 525-line television standards, the source coder operates on pictures based on a common intermediate format (CIF). The standards of the input and output television signals, which may, for example, be composite or component, analogue or digital and the methods of performing any necessary conversion to and from the source coding format are not subject to recommendation.

2.2 Digital output and input

The video coder provides a self-contained digital bit stream which may be combined with other multi-facility signals (for example as defined in Recommendation H.221). The video decoder performs the reverse process.

2.3 Sampling frequency

Pictures are sampled at an integer multiple of the video line rate. This sampling clock and the digital network clock are asynchronous.

2.4 Source coding algorithm

A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The decoder has motion compensation capability, allowing optional incorporation of this technique in the coder.

2.5 Bit rate

This Recommendation is primarily intended for use at video bit rates between approximately 40 kbit/s and 2 Mbit/s.

2.6 Symmetry of transmission

The codec may be used for bidirectional or unidirectional visual communication.

2.7 Error handling

The transmitted bit-stream contains a BCH (511,493) Forward Error Correction Code. Use of this by the decoder is optional.

2.8 Multipoint operation

Features necessary to support switched multipoint operation are included.

3. Source coder

3.1 Source format

The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is ±50 ppm.

Pictures are coded as luminance and two colour difference components (Y, C_B and C_R). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

Black = 16
White = 235
Zero colour difference = 128
Peak colour difference = 16 and 240

These values are nominal ones and the coding algorithm functions with input values of 1 through to 254.

Two picture scanning formats are specified.

In the first format (CIF), the luminance sampling structure is 352 pels per line, 288 lines per picture in an orthogonal arrangement. Sampling of each of the two colour difference components is at 144 lines, 176 pels per line, orthogonal. Colour difference samples are sited such that their clock boundaries coincide with luminance block boundaries as shown in Figure 2/H.261. The picture area covered by these numbers of pels and lines has an aspect ratio of 4:3 and corresponds to the active portion of the local standard video input.

Note - The number of pels per line is compatible with sampling the active portions of the luminance and colour difference signals from 525- or 625-line sources at 6.75 and 3.375 MHz respectively. These frequencies have a simple relationship to those in CCIR Recommendation 601.

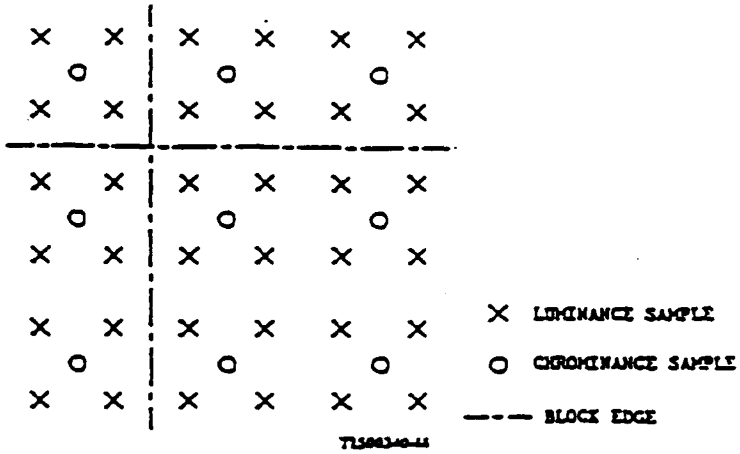


FIGURE 2/H.261

Positioning of luminance and chrominance samples

The second format, Quarter-CIF (QCIF), has half the number of pels and half the number of lines stated above. All codecs must be able to operate using QCIF. Some codecs can also operate with CIF.

Means shall be provided to restrict the maximum picture rate of encoders by having at least 0, 1, 2 or 3 non-transmitted pictures between transmitted ones. Selection of this minimum number and CIF or QCIF shall be by external means (for example via Recommendation H.221).

3.2 Video source coding algorithm

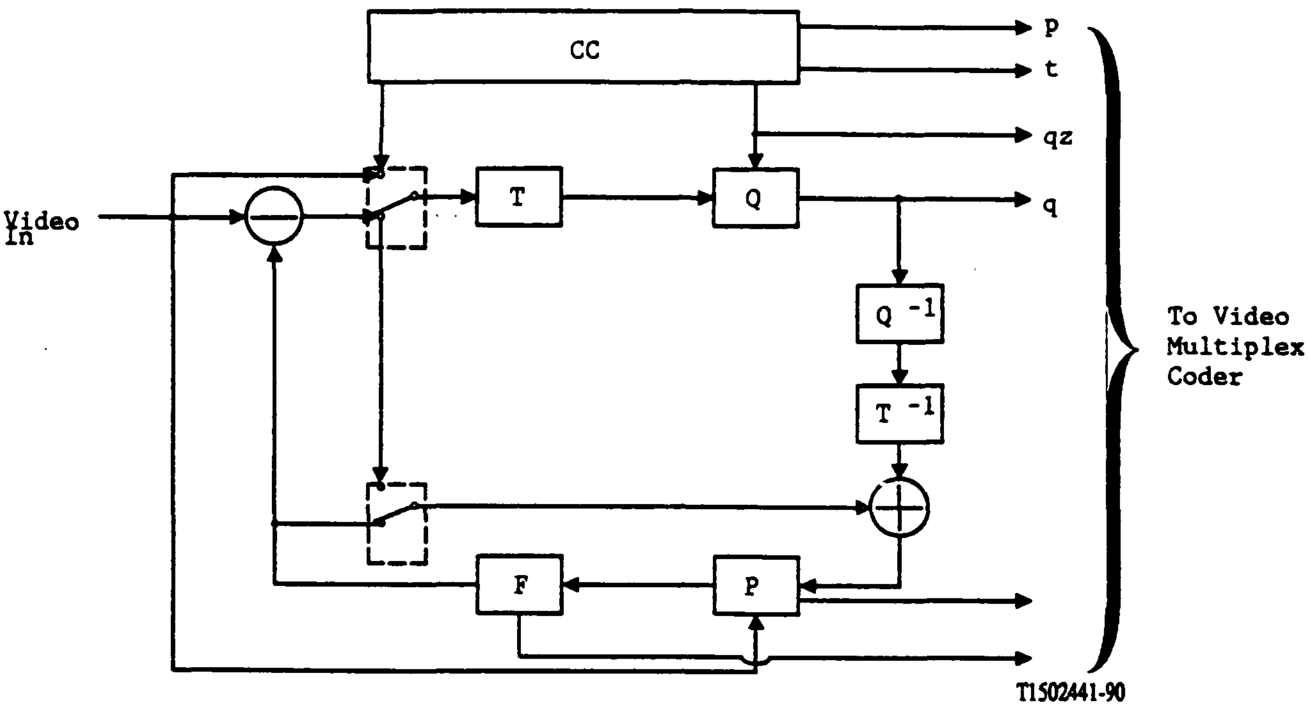
The source coder is shown in generalized form in Figure 3/H.261. The main elements are prediction, block transformation and quantization.

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. Further, four luminance blocks and the two spatially corresponding colour difference blocks are combined to form a macroblock as shown in Figure 10/H.261 of § 4.2.4.

The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the coding control strategy. Transmitted blocks are transformed and resulting coefficients are quantized and variable length coded.

3.2.1 Prediction

The prediction is inter-picture and may be augmented by motion compensation (§ 3.2.2) and a spatial filter (§ 3.2.3).



T: Transform
Q: Quantizer
P: Picture Memory with motion compensated variable delay
F: Loop filter
CC: Coding control

p: Flag for INTRA/INTER
t: Flag for transmitted or not
qz: Quantizer indication
q: Quantizing index for transform coefficients
v: Motion vector
f: Switching on/off of the loop filter

FIGURE 3/H.261

Source coder

3.2.2 Motion compensation

Motion compensation (MC) is optional in the encoder. The decoder will accept one vector per macroblock. Both horizontal and vertical components of these motion vectors have integer values not exceeding ±15. The vector is used for all four luminance blocks in the macroblock. The motion vector for both colour difference blocks is derived by halving the component values of the macroblock vector and truncating the magnitude parts towards zero to yield integer components.

A positive value of the horizontal or vertical component of the motion vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right or below the pels being predicted.

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

3.2.3 Loop filter

The prediction process may be modified by a two-dimensional spatial filter (FIL) which operates on pels within a predicted 8 by 8 block.

The filter is separable into one-dimensional horizontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4 except at block edges where one of the taps would fall outside the block. In such cases the 1-D filter is changed to have coefficients of 0, 1, 0. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter is switched on/off for all six blocks in a macroblock according to the macroblock type (see § 4.2.3 MTYPE).

3.2.4 Transformer

Transmitted blocks are first processed by a separable two-dimensional Discrete Cosine Transform of size 8 by 8. The output from the inverse transform ranges from -256 to +255 after clipping to be represented with 9 bits. The transfer function of the inverse transform is given by:

$$f(x,y) = 1/4 \sum_{u=0}^{7} \sum_{v=0}^{7} C(u)\, C(v)\, F(u,v) \cos[P(2x+1)u/16] \cos[P(2y+1)v/16]$$

with u, v, x, y = 0, 1, 2, ...7

where x,y = spatial coordinates in the pel domain

u,v = coordinates in the transform domain

$C(u) = 1/\sqrt{2}$ for u = 0, otherwise 1
$C(v) = 1/\sqrt{2}$ for v = 0, otherwise 1

Note - Within the block being transformed, x = 0 and y = 0 refer to the pel nearest the left and top edges of the picture respectively.

The arithmetic procedures for computing the transforms are not defined, but the inverse one should meet the error tolerance specified in Annex 1.

3.2.5 Quantization

The number of quantizers is 1 for the INTRA dc coefficient and 31 for all other coefficients. Within a macroblock the same quantizer is used for all coefficients except the INTRA dc one. The decision levels are not defined. The INTRA dc coefficient is nominally the transform value linearly quantized with a stepsize of 8 and no dead-zone. Each of the other 31 quantizers is also nominally linear but with a central dead-zone around zero and with a step size of an even value in the range 2 to 62.

The reconstruction levels are as defined in § 4.2.4.

Note - For the smaller quantization step sizes, the full dynamic range of the transform coefficients cannot be represented.

3.2.6 Clipping of reconstructed picture

To prevent quantization distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. The clipping function is applied to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

3.3 Coding control

Several parameters may be varied to control the rate of generation of coded video data. These include processing prior to the source coder, the quantizer, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures.

3.4 Forced updating

This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update pattern is not defined. For control of accumulation of inverse transform mismatch error a macroblock should be forcibly updated at least once per every 132 times it is transmitted.

4. Video multiplex coder

4.1 Data structure

Unless specified otherwise the most significant bit is transmitted first. This is bit 1 and is the leftmost bit in the code tables in this document. Unless specified otherwise all unused or spare bits are set to "1". Spare bits must not be used until their functions are specified by the CCITT.

4.2 Video multiplex arrangement

The video multiplex is arranged in a hierarchical structure with four layers. From top to bottom the layers are:

Picture
Group of blocks (GOB)
Macroblock (MB)
Block

A syntax diagram of the video multiplex coder is shown in Figure 4/H.261. Abbreviations are defined in later sections.

4.2.1 Picture layer

Data for each picture consists of a picture header followed by data for GOBs. The structure is shown in Figure 5/H.261. Picture headers for dropped pictures are not transmitted.

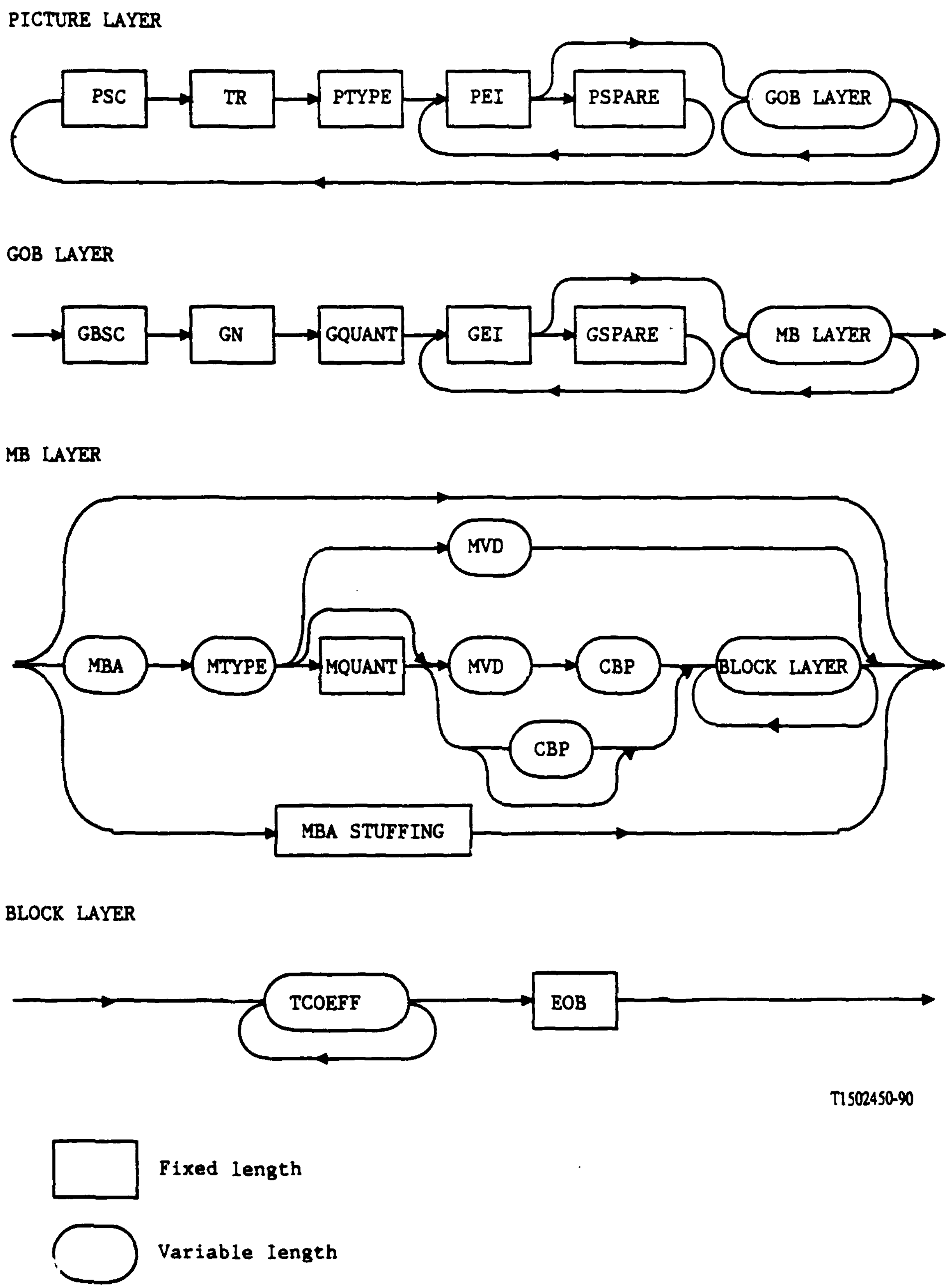


FIGURE 4/H.261

<u>Syntax diagram for the video multiplex coder</u>

```
 ------------------------   ----------------   -----------
| PSC : TR : PTYPE : PEI :    PSPASRE : PEI :     GOB Data |
 ------------------------   ----------------   -----------
```

FIGURE 5/H.261

Structure of picture layer

Picture Start Code (PSC) 20 bits

A word of 20 bits. Its value is 0000 0000 0000 0001 0000.

Temporal Reference (TR) 5 bits

A 5-bit number which can have 32 possible values. It is formed by incrementing its value in the previously transmitted picture header by one plus the number of non-transmitted pictures (at 29.97 Hz) since that last transmitted one. The arithmetic is performed with only the five LSBs.

Type Information (PTYPE) 6 bits

Information about the complete picture:

- Bit 1: Split screen indicator. "0" off, "1" on.
- Bit 2: Document camera indicator. "0" off, "1" on.
- Bit 3: Freeze Picture Release. "0" off, "1" on.
- Bit 4: Source Format. "0" QCIF, "1" CIF.
- Bits 5 to 6: Spare.

Extra Insertion Information (PEI) 1 bit

A bit which when set to "1" signals the presence of the following optional data field.

Spare Information (PSPARE) 0/8/16 ... bits

If PEI is set to "1", then 9 bits follow consisting of 8 bits of data (PSPARE) and then another PEI bit to indicate if a further 9 bits follow and so on. Encoders must not insert PSPARE until specified by the CCITT. Decoders must be designed to discard PSPARE if PEI is set to 1. This will allow the CCITT to specify future "backward" compatible additions in PSPARE.

4.2.2 Group of blocks layer

Each picture is divided into groups of blocks (GOBs). A group of blocks (GOB) comprises one twelfth of the CIF or one third of the QCIF picture areas (see Figure 6/H.261). A GOB relates to 176 pels by 48 lines of Y and the spatially corresponding 88 pels by 24 lines of each of C_B and C_R.

Data for each group of blocks consists of a GOB header followed by data for macroblocks. The structure is shown in Figure 7/H.261. Each GOB header is transmitted once between Picture Start Codes in the CIF or QCIF sequence numbered in Figure 6/H.261, even if no macroblock data is present in that GOB.

Group of blocks Start Code (GBSC) 16 bits

A word of 16 bits, 0000 0000 0000 0001.

1	2
3	4
5	6
7	8
9	10
11	12

CIF

1
3
5

QCIF

FIGURE 6/H.261

<u>Arrangement of GOBs in a picture</u>

GBSC	GN	GQUANT	GEI	GSPARE	GEI	MB Data

FIGURE 7/H.261

<u>Structure of group of blocks layer</u>

Group Number (GN)   4 bits

Four bits indicating the position of the group of blocks. The bits are the binary representation of the number in Figure 6/H.261. Group numbers 13, 14 and 15 are reserved for future use. Group number 0 is used in the PSC.

Quantizer Information (GQUANT)   5 bits

A fixed length codeword of 5 bits which indicates the quantizer to be used in the group of blocks until overridden by any subsequent MQUANT. The codewords are the natural binary representations of the values of QUANT (§ 4.2.4) which, being half the step sizes, range from 1 to 31.

Extra Insertion Information (GEI) 1 bit

A bit which when set to "1" signals the presence of the following optional data field.

Spare Information (GSPARE)   0/8/16 ... bits

If GEI is set to "1", then 9 bits follow consisting of 8 bits of data (GSPARE) and then another GEI bit to indicate if a further 9 bits follow and so on. Encoders must not insert GSPARE until specified by the CCITT. Decoders must be designed to discard GSPARE if GEI is set to 1. This will allow the CCITT to specify future "backward" compatible additions in GSPARE.

<u>Note</u> - Emulation of start codes may occur if the future specification of GSPARE has no restrictions on the final GSPARE data bits.

4.2.3 Macroblock layer

Each GOB is divided into 33 macroblocks as shown in Figure 8/H.261. A macroblock relates to 16 pels by 16 lines of Y and the spatially corresponding 8 pels by 8 lines of each of C_B and C_R.

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

FIGURE 8/H.261

Arrangement of macroblocks in a GOB

Data for a macroblock consists of a MB Header followed by data for blocks (Figure 9/H.261). MQUANT, MVD and CBP are present when indicated by MTYPE.

MBA	MTYPE	MQUANT	MVD	CBP	Block Data

FIGURE 9/H.261

Structure of macroblock layer

Macroblock Address (MBA)   Variable Length

A variable length codeword indicating the position of a macroblock within a group of blocks. The transmission order is as shown in Figure 8/H.261. For the first transmitted macroblock in a GOB, MBA is the absolute address in Figure 8/H.261. For subsequent macroblocks, MBA is the difference between the absolute addresses of the macroblock and the last transmitted macroblock. The code table for MBA is given in Table 1/H.261.

An extra codeword is available in the table for bit stuffing immediately after a GOB header or a coded macroblock (MBA Stuffing). This codeword should be discarded by decoders.

The VLC for start code is also shown in Table 1/H.261.

TABLE 1/H.261

VLC table for macroblock addressing

MBA	CODE	MBA	CODE
1	1	17	0000 0101 10
2	011	18	0000 0101 01
3	010	19	0000 0101 00
4	0011	20	0000 0100 11
5	0010	21	0000 0100 10
6	0001 1	22	0000 0100 011
7	0001 0	23	0000 0100 010
8	0000 111	24	0000 0100 001
9	0000 110	25	0000 0100 000
10	0000 1011	26	0000 0011 111
11	0000 1010	27	0000 0011 110
12	0000 1001	28	0000 0011 101
13	0000 1000	29	0000 0011 100
14	0000 0111	30	0000 0011 011
15	0000 0110	31	0000 0011 010
16	0000 0101 11	32	0000 0011 001
		33	0000 0011 000
		MBA Stuffing	0000 0001 111
		Start code	0000 0000 0000 0001

MBA is always included in transmitted macroblocks.

Macroblocks are not transmitted when the contain no information for that part of the picture.

Type Information (MTYPE)   Variable Length

Variable length codewords giving information about the macroblock and which data elements are present. Macroblock types, included elements and VLC words are listed in Table 2/H.261.

MTYPE is always included in transmitted macroblocks.

TABLE 2/H.261

VLC table for MTYPE

Prediction	MQUANT	MVD	CBP	TCOEFF	VLC
Intra				x	0001
Intra	x			x	0000 001
Inter			x	x	1
Inter	x		x	x	0000 1
Inter + MC		x			0000 0000 1
Inter + MC		x	x	x	0000 0001
Inter + MC	x	x	x	x	0000 0000 01
Inter + MC + FIL		x			001
Inter + MC + FIL		x	x	x	01
Inter + MC + FIL	x	x	x	x	0000 01

Note 1 - "x" means that the item is present in the macroblock.

Note 2 - It is possible to apply the filter in a non-motion compensated macroblock by declaring it as MC + FIL but with a zero vector.

Quantizer (MQUANT) 5 bits

MQUANT is present only if so indicated by MTYPE.

A codeword of 5 bits signifying the quantizer to be used for this and any following blocks in the group of blocks until overridden by any subsequent MQUANT.

Codewords for MQUANT are the same as for GQUANT.

Motion Vector Data (MVD) Variable length

Motion Vector Data is included for all MC macroblocks. MVD is obtained from the macroblock vector by subtracting the vector of the preceding macroblock. For this calculation the vector of the preceding macroblock is regarded as zero in the following three situations:

1) Evaluating MVD for macroblocks 1, 12 and 23.

2) Evaluating MVD for macroblocks in which MBA does not represent a difference of 1.

3) MTYPE of the previous macroblock was not MC.

MVD consists of a variable length codeword for the horizontal component followed by a variable length codeword for the vertical component. Variable length codes are given in Table 3/H.261.

Advantage is taken of the fact that the range of motion vector values is constrained. Each VLC word represents a pair of difference values. Only one of the pair will yield a macroblock vector falling within the permitted range.

Coded Block Pattern (CBP)   Variable length

CBP is present if indicated by MTYPE. The codeword gives a pattern number signifying those blocks in the macroblock for which at least one transform coefficient is transmitted. The pattern number is given by:

$$32*P_1 + 16*P_2 + 8*P_3 + 4*P_4 + 2*P_5 + P_6$$

where P_n is 1 if any coefficient is present for block n, else 0. Block numbering is given in Figure 10/H.261.

The codewords for CBP are given in Table 4/H.261.

TABLE 3/H.261

VLC table for MVD

MVD	CODE
-16 & 16	0000 0011 001
-15 & 17	0000 0011 011
-14 & 18	0000 0011 101
-13 & 19	0000 0011 111
-12 & 20	0000 0100 001
-11 & 21	0000 0100 011
-10 & 22	0000 0100 11
-9 & 23	0000 0101 01
-8 & 24	0000 0101 11
-7 & 25	0000 0111
-6 & 26	0000 1001
-5 & 27	0000 1011
-4 & 28	0000 111
-3 & 29	0001 1
-2 & 30	0011
-1	011
0	1
1	010
2 & -30	0010
3 & -29	0001 0
4 & -28	0000 110
5 & -27	0000 1010
6 & -26	0000 1000
7 & -25	0000 0110
8 & -24	0000 0101 10
9 & -23	0000 0101 00
10 & -22	0000 0100 10
11 & -21	0000 0100 010
12 & -20	0000 0100 000
13 & -19	0000 0011 110
14 & -18	0000 0011 100
15 & -17	0000 0011 010

TABLE 4/H.261

VLC table for CBP

CBP	CODE	CBP	CODE
60	111	35	0001 1100
4	1101	13	0001 1011
8	1100	49	0001 1010
16	1011	21	0001 1001
32	1010	41	0001 1000
12	1001 1	14	0001 0111
48	1001 0	50	0001 0110
20	1000 1	22	0001 0101
40	1000 0	42	0001 0100
28	0111 1	15	0001 0011
44	0111 0	51	0001 0010
52	0110 1	23	0001 0001
56	0110 0	43	0001 0000
1	0101 1	25	0000 1111
61	0101 0	37	0000 1110
2	0100 1	26	0000 1101
62	0100 0	38	0000 1100
24	0011 11	29	0000 1011
36	0011 10	45	0000 1010
3	0011 01	53	0000 1001
63	0011 00	57	0000 1000
5	0010 111	30	0000 0111
9	0010 110	46	0000 0110
17	0010 101	54	0000 0101
33	0010 100	58	0000 0100
6	0010 011	31	0000 0011 1
10	0010 010	47	0000 0011 0
18	0010 001	55	0000 0010 1
34	0010 000	59	0000 0010 0
7	0001 1111	27	0000 0001 1
11	0001 1110	39	0000 0001 0
19	0001 1101		

4.2.4 Block layer

A macroblock comprises four luminance blocks and one of each of the two colour difference blocks (Figure 10/H.261).

```
 -------        -------        -------
| 1 | 2 |      |       |      |       |
|-------|      |   5   |      |   6   |
| 3 | 4 |      |       |      |       |
 -------        -------        -------
    Y              CB             CR
```

FIGURE 10/H.261

Arrangement of blocks in a macroblock

Data for a block consists of codewords for transform coefficients followed by an end of block marker (Figure 11/H.261). The order of block transmission is as in Figure 10/H.261.

```
 -------------
| TCOEFF | EOB |
 -------------
```

FIGURE 11/H.261

Structure of block layer

Transform Coefficients (TCOEFF)

Transform coefficient data is always present for all six blocks in a macroblock when MTYPE indicates INTRA. In other cases MTYPE and CBP signal which blocks have coefficient data transmitted for them. The quantized transform coefficients are sequentially transmitted according to the sequence given in Figure 12/H.261.

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

```
---------> increasing cycles per
|          picture width
|
|
|
|
V

increasing cycles per
picture height
```

FIGURE 12/H.261

Transmission order for transform coefficients

The most commonly occurring combinations of successive zeros (RUN) and the following value (LEVEL) are encoded with variable length codes. Other combinations of (RUN, LEVEL) are encoded with a 20-bit word consisting of 6 bits ESCAPE, 6 bits RUN and 8 bits LEVEL. For the variable length encoding there are two code tables, one being used for the first transmitted LEVEL in INTER, INTER+MC and INTER+MC+FIL blocks, the second for all other LEVELs except the first one in INTRA blocks which is fixed length coded with 8 bits.

Codes are given in Table 5/H.261.

TABLE 5/H.261

<u>VLC table for TCOEFF</u>

The most commonly occurring combinations of zero-run and the following value are encoded with variable length codes as listed in the table below. End of block (EOB) is in this set. Because CBP indicates those blocks with no coefficient data, EOB cannot occur as the first coefficient. Hence EOB can be removed from the VLC table for the first coefficient.

The last bit "s" denotes the sign of the level, "0" for positive
"1" for negative.

RUN	| LEVEL |	CODE
EOB		10
0	1	1s IF FIRST COEFFICIENT IN BLOCK (<u>Note</u> - Never used in INTRA macroblocks)
0	1	11s NOT FIRST COEFFICIENT IN BLOCK
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1010 1s
2	1	0101 s
2	2	0000 100s
2	3	0000 0010 11s
2	4	0000 0001 0100 s
2	5	0000 0000 1010 0s

3	1	0011 1s
3	2	0010 0100 s
3	3	0000 0001 1100 s
3	4	0000 0000 1001 1s
4	1	0011 0s
4	2	0000 0011 11s
4	3	0000 0001 0010 s
5	1	0001 11s
5	2	0000 0010 01s
5	3	0000 0000 1001 0s
6	1	0001 01s
6	2	0000 0001 1110 s
7	1	0001 00s
7	2	0000 0001 0101 s
8	1	0000 111s
8	2	0000 0001 0001 s
9	1	0000 101s
9	2	0000 0000 1000 1s
10	1	0010 0111 s
10	2	0000 0000 1000 0s
11	1	0010 0011 s
12	1	0010 0010 s
13	1	0010 0000 s
14	1	0000 0011 10s
15	1	0000 0011 01s
16	1	0000 0010 00s
17	1	0000 0001 1111 s
18	1	0000 0001 1010 s
19	1	0000 0001 1001 s
20	1	0000 0001 0111 s
21	1	0000 0001 0110 s
22	1	0000 0000 1111 1s
23	1	0000 0000 1111 0s
24	1	0000 0000 1110 1s
25	1	0000 0000 1110 0s
26	1	0000 0000 1101 1s
ESCAPE		0000 01

The remaining combinations of (RUN, LEVEL) are encoded with a 20-bit word[1] consisting of 6 bits ESCAPE, 6 bits RUN and 8 bits LEVEL.

[1] Use of this 20-bit word form for encoding the combinations listed in the VLC table is not prohibited.

RUN is a 6 bit fixed length code. LEVEL is an 8 bit fixed length code.

RUN	CODE
0	0000 00
1	0000 01
2	0000 10
.	.
.	.
63	1111 11

LEVEL	CODE
-128	FORBIDDEN
-127	1000 0001
.	.
-2	1111 1110
-1	1111 1111
0	FORBIDDEN
1	0000 0001
2	0000 0010
.	.
127	0111 1111

For all coefficients other than the INTRA dc one the reconstruction levels (REC) are in the range -2048 to 2047 and are given by clipping the results of the following formulae:

$$\left.\begin{array}{ll} REC = QUANT*(2*LEVEL+1) & ;\ LEVEL>0 \\ REC = QUANT*(2*LEVEL-1) & ;\ LEVEL<0 \end{array}\right\} QUANT = \text{"odd"}$$

$$\left.\begin{array}{ll} REC = QUANT*(2*LEVEL+1)-1; & LEVEL>0 \\ REC = QUANT*(2*LEVEL-1)+1; & LEVEL<0 \end{array}\right\} QUANT = \text{"even"}$$

$$REC = 0;\ LEVEL=0$$

<u>Note</u> - QUANT ranges from 1 to 31 and is transmitted by either GQUANT OR MQUANT.

TABLE 6/H.261

<u>Reconstruction levels (REC)</u>

LEVEL	QUANT 1	2	3	4	.	8	9	.	17	18	.	30	31
-127	-255	-509	-765	-1019	.	-2039	-2048	.	-2048	-2048	.	-2048	-2048
-126	-253	-505	-759	-1011	.	-2023	-2048	.	-2048	-2048	.	-2048	-2048
.	.	.	.	.	.	.	.	.	.	.	.	.	.
-2	-5	-9	-15	-19	.	-39	-45	.	-85	-89	.	-149	-155
-1	-3	-5	-9	-11	.	-23	-27	.	-51	-53	.	-89	-93
0	0	0	0	0	.	0	0	.	0	0	.	0	0
1	3	5	9	11	.	23	27	.	51	53	.	89	93
2	5	9	15	19	.	39	45	.	85	89	.	149	155
3	7	13	21	27	.	55	63	.	119	125	.	209	217
4	9	17	27	35	.	71	81	.	153	161	.	269	279
5	11	21	33	43	.	87	99	.	187	197	.	329	341
.	.	.	.	.	.	.	.	.	.	.	.	.	.
56	113	225	339	451	.	903	1017	.	1921	2033	.	2047	2047
57	115	229	345	459	.	919	1035	.	1955	2047	.	2047	2047
58	117	233	351	467	.	935	1053	.	1989	2047	.	2047	2047
59	119	237	357	475	.	951	1071	.	2023	2047	.	2047	2047
60	121	241	363	483	.	967	1089	.	2047	2047	.	2047	2047
.	.	.	.	.	.	.	.	.	.	.	.	.	.
125	251	501	753	1003	.	2007	2047	.	2047	2047	.	2047	2047
126	253	505	759	1011	.	2023	2047	.	2047	2047	.	2047	2047
127	255	509	765	1019	.	2039	2047	.	2047	2047	.	2047	2047

Note - Reconstruction levels are symmetrical with respect to the sign of LEVEL except for 2047/-2048.

For INTRA blocks the first coefficient is nominally the transform dc value linearly quantized with a step size of 8 and no dead-zone. The resulting values are represented with 8 bits. A nominally black block will give 0001 0000 and a nominally white one 1110 1011. The code 0000 0000 is not used. The code 1000 0000 is not used, the reconstruction level of 1024 being coded as 1111 1111 (see Table 7/H.261).

Coefficients after the last non-zero one are not transmitted. EOB (end of block code) is always the last item in blocks for which coefficients are transmitted.

TABLE 7/H.261

Reconstruction levels for INTRA-mode dc coefficient

FLC	Reconstruction level into inverse transform
0000 0001 (1)	8
0000 0010 (2)	16
0000 0011 (3)	24
.	.
.	.
0111 1111 (127)	1016
1111 1111 (255)	1024
1000 0001 (129)	1032
.	.
.	.
1111 1101 (253)	2024
1111 1110 (254)	2032

Note - The decoded value corresponding to FLC "n" is 8n except FLC 255 gives 1024.

4.3 Multipoint considerations

The following facilities are provided to support switched multipoint operation.

4.3.1 Freeze picture request

Causes the decoded to freeze its displayed picture until a freeze picture release signal is received or a timeout period of at least six seconds has expired. The transmission of this signal is via external means (for example by H.221).

4.3.2 Fast update request

Causes the encoder to encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this signal is via external means (for example by H.221).

4.3.3 Freeze picture release

A signal from an encoder which has responded to a Fast Update Request and allows a decoder to exit from its freeze picture mode and display decoded pictures in the normal manner. This signal is transmitted by bit 3 of PTYPE (see § 4.2.1) in the picture header of the first picture coded in response to the Fast Update Request.

5. Transmission coder

5.1 Bit rate

The transmission clock is provided externally (for example from an I.420 interface).

5.2 Video data buffering

The encoder must control its output bitstream to comply with the requirements of the Hypothetical Reference Decoder defined in Annex 2.

When operating with CIF the number of bits created by coding any single picture must not exceed 256 kbit/s. K = 1024.

When operating with QCIF the number of bits created by coding any single picture must not exceed 64 kbit/s.

In both the above cases the bit count includes the Picture Start Code and all other data related to that picture including PSPARE, GSPARE and MBA Stuffing. The bit count does not include error correction framing bits, fill indicator (Fi), fill bits or error correction parity information described in § 5.4 below.

Video data must be provided on every valid clock cycle. This can be ensured by the use of either the fill bit indicator (Fi) and subsequent fill all 1's bits in the error corrector block framing (see Figure 13/H.261) or MBA Stuffing (§ 4.2.3) or both.

5.3 Video coding delay

This item is included in this Recommendation because the video encoder and video decoder delays need to be known to allow audio compensation delays to be fixed when H.261 is used to form part of a conversational service. This will allow lip synchronization to be maintained. Annex 3 recommends a method by which the delay figures are established. Other delay measurement methods may be used but they must be designed in a way to produce similar results to the method given in Annex 3.

5.4 Forward Error Correction for coded video signal

5.4.1 Error correcting code

The transmitted bitstream contains a BCH (511,493) Forward Error Correction Code. Use of this by the decoder is optional.

5.4.2 Generator polynomial

$$g(x) = (x^9 + x^4 + 1)(x^9 + x^6 + x^4 + x^3 + 1)$$

Example: for the input data of "01111 ... 11" (493 bits) the resulting correction parity bits are "011011010100011011" (18 bits).

5.4.3 Error correction framing

To allow the video data and error correction parity information to be identified by a decoder an error correction framing pattern is included. This consists of a multiframe of eight frames, each frame comprising 1 bit framing, 1 bit fill indicator (Fi), 492 bits of coded data (or fill all 1s) and 18 bits parity. The frame alignment pattern is:

$$(S_1S_2S_3S_4S_5S_6S_7S_8) = (00011011).$$

See Figure 13/H.261 for the frame arrangement. The parity is calculated against the 493-bits including fill indicator (Fi).

The fill indicator (Fi) can be set to zero by an encoder. In this case only 492 consecutive fill bits (fill all 1s) plus parity are sent and no coded data is transmitted. This may be used to meet the requirement in § 5.2 to provide video data on every valid clock cycle.

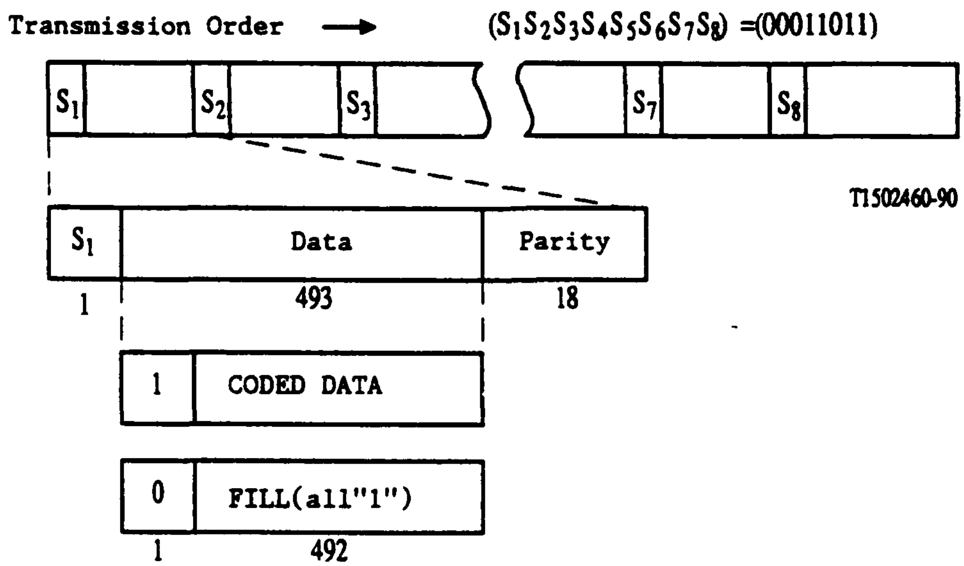


FIGURE 13/H.261

Error correcting frame

5.4.4 Relock time for error corrector framing

Three consecutive error correction framing sequences (24 bits) should be received before frame lock is deemed to have been achieved. The decoder should be designed such that frame lock will be re-established within 34000 bits after an error corrector framing phase change.

Note - This assumes that the video data does not contain three correctly phased emulations of the error correction framing sequence during the relocking period.

ANNEX 1

(to Recommendation H.261)

Inverse transform accuracy specification

1. Generate random integer pel data values in the range -L to +H according to the random number generator given below ("C" version). Arrange into 8 by 8 blocks. Data set of 10,000 blocks should each be generated for (L = 256, H = 255), (L = H = 5) and (L = H = 300).

2. For each 8 by 8 block, perform a separable, orthonormal, matrix multiply, Forward Discrete Cosine Transform using at least 64-bit floating point accuracy.

$$F(u,v) = 1/4\ C(u)\ C(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos[Pi(2x + 1)u/16] \cos[Pi(2y + 1)v/16]$$

with u, v, x, y = 0, 1, 2, ...7

where x,y = spatial coordinates in the pel domain

u,v = coordinates in the transform domain

$C(u) = 1/\sqrt{2}$ for u = 0, otherwise 1
$C(v) = 1/\sqrt{2}$ for v = 0, otherwise 1

3. For each block, round the 64 resulting transformed coefficients to the nearest integer values. Then clip them to the range -2048 to +2047. This is the 12-bit input data to the inverse transform.

4. For each 8 by 8 block of 12-bit data produced by step 3, perform a separable, orthonormal, matrix multiply, Inverse Discrete Transform (IDCT) using at least 64-bit floating point accuracy. Round the resulting pels to the nearest integer and clip to the range -256 to +255. These blocks of 8 by 8 pels are the "reference" IDCT input data.

5. For each 8 by 8 block produced by step 3, apply the IDCT under test and clip the output to the range -256 to +255. These blocks of 8 by 8 pels are the "test" IDCT output data.

6. For each of the 64 IDCT output pels, and for each of the 10,000 block data sets generated above, measure the peak, mean and mean square error between the "reference" and the "test" data.

7. For any pel, the peak error should not exceed 1 in magnitude.

For any pel, the mean square error should not exceed 0.06.

Overall, the mean square error should not exceed 0.02.

For any pel, the mean error should not exceed 0.015 in magnitude.

Overall, the mean error should not exceed 0.0015 in magnitude.

8. All zeros in must produce all zeros out.

9. Re-run the measurements using exactly the same data values of step 1, but change the sign on each pel.

"C" Program for random number generation

```
/* L and H must be long, that is 32 bits */
long rand(L,H)
long     L,H;
{
  static long  randx = 1;  /* long is 32 bits */
  static double  z = (double)0x7fffffff;

  long i,j;
  double x;                  /* double is 64 bits */

  randx = (randx * 1103515245) + 12345;
  i = randx & 0x7ffffffe;      /* keep 30 bits */
  x = ( (double)i ) / z;       /* range 0 to 0.99999... */
  x * = (L+H+1);               /* range 0 to < L+H+1 */
  j = x;                       /* truncate to integer */
  return( j - L );             /* range -L to H */
}
```

ANNEX 2

(to Recommendation H.261)

Hypothetical Reference Decoder

The Hypothetical Reference Decoder (HRD) is defined as follows:

1. The HRD and the encoder have the same clock frequency as well as the same CIF rate, and are operated synchronously.

2. The HRD receiving buffer size is (B + 256 kbit/s). The value of B is defined as follows:

$B = 4R_{max}/29.97$ where R_{max} is the maximum video bit rate to be used in the connection.

3. The HRD buffer is initially empty.

4. The HRD buffer is examined at CIF intervals (≃33 ms). If at least one complete coded picture is in the buffer then all the data for the earliest picture is instantaneously removed (e.g. at t_{n+1} in Figure A.1/H.261). Immediately after removing the above data the buffer occupancy must be less than B. This is a requirement on the coder output bitstream including coded picture data and MBA stuffing but not error correction framing bits, fill indicator (Fi), fill bits or error correction parity information described in § 5.4.

To meet this requirement the number of bits for the (N+1)th coded picture d_{N+1} must satisfy:

$$d_{N+1} > B_N + \int_{t_N}^{t_{N+1}} R(t)dt - B$$

where B_N is buffer occupancy just after the time t_N,

t_N is the time the Nth coded picture is removed from the HRD buffer,

$R(t)$ is the video bit rate at the time t.

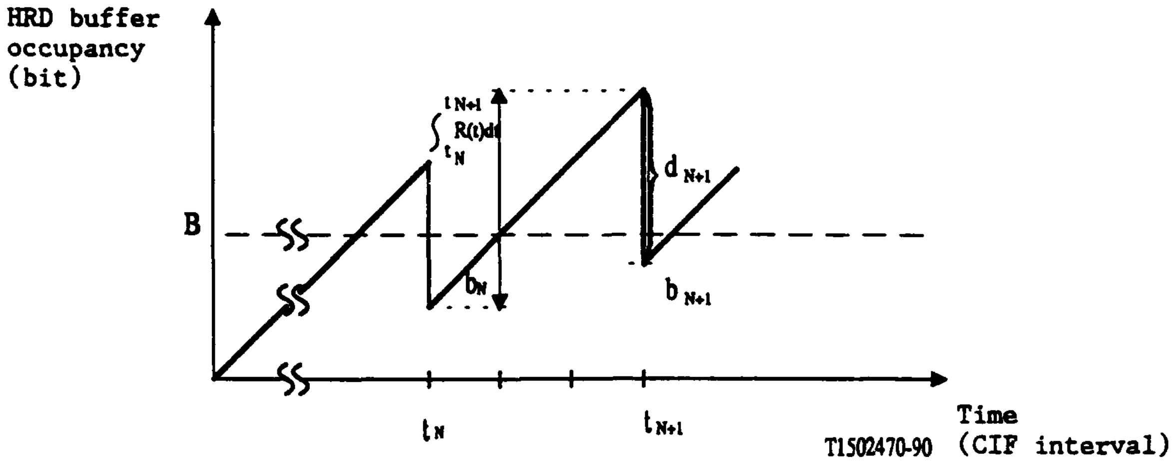


FIGURE A.1/H.261

HRD buffer occupancy

Note - Time (t_{N+1} - t_N) is an integer number of CIF picture periods (1/29.97, 2/29.97, 3/29.97, ...).

ANNEX 3

(to Recommendation H.261)

Codec delay measurement method

The video encoder and video decoder delays will vary depending on implementation. The delay will also depend on the picture format (QCIF, CIF) and data rate in use. This section specifies the method by which the delay figures are established for a particular design. To allow correct audio delay compensation the overall video delay needs to be established from a user perception point of view under typical viewing conditions.

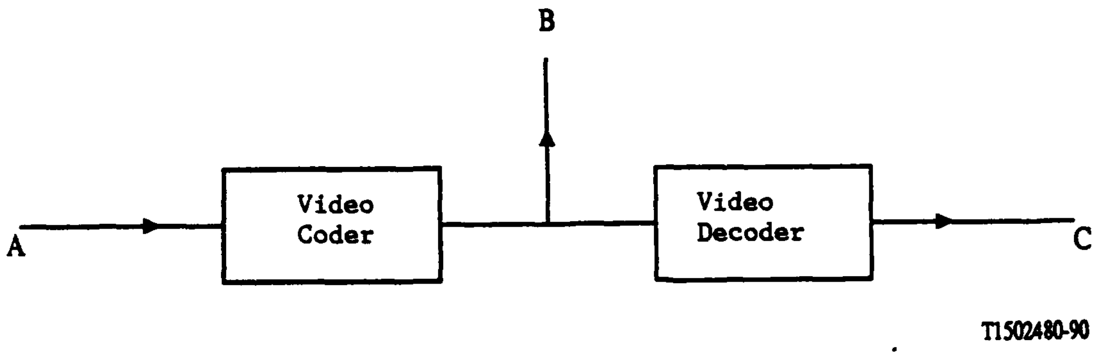


FIGURE A.2/H.261

Measuring points

Point A is the video input to the video coder. Point B is the channel output from the video terminal (i.e. including any FEC, channel framing etc.). Point C is the video output from the decoder.

A video sequence lasting more than 100 seconds is connected to the video coder input (point A) in Figure A.2/H.261 above. The video sequence should have the following characteristics:

- it should contain a typical moving scene consistent with the intended purpose of the video codec;
- it should produce a minimum coded picture rate of 7.5 Hz at the bit rate in use;
- it should contain a visible identification mark at intervals throughout the length of the sequence. The visible identification should change every 97 video input frames and be located within the picture area represented by the first GOB in the picture. For example, the first block in the picture could change from black to white at intervals of 97 video frame periods. The identification mark should be chosen so that it can be detected at point B and does not significantly contribute to the overall coding performance.

The codec and video sequence should be arranged so that the bitstream contains less than 10% stuffing (MBA stuffing + error correction fill bits).

The encoder delay is obtained by measuring the time from when the visible identification changes at point A to the time that the change is detected at point B. Similarly, the decoder delay is obtained by taking measurements at points B and C.

Several measurements should be made during the sequence length and the average period obtained. Several tests should be made to ensure that a consistent average figure can be obtained for both encoder and decoder delay times.

Average results should be obtained for each combination of picture format and bit rate within the capability of the particular codec design.

Note - Due to pre- and post-temporal processing it may be necessary to take a mid-level for establishing the transition of the identification mark at points B and C.

5. Recommendation H.320

NARROW-BAND VISUAL TELEPHONE SYSTEMS AND TERMINAL EQUIPMENT

CONTENTS

1. Scope

This Recommendation covers the technical requirements for narrow-band visual telephone services defined in H.200/AV.120-Series Recommendations, where channel rates do not exceed 1920 kbit/s.

Note - It is anticipated that this Recommendation will be extended to a number of Recommendations each of which would cover a single videoconferencing or videophone service (narrow-band, broadband ...). However, large parts of these Recommendations would have identical wording, while in the points of divergence the actual choices between alternatives have not yet been made; for the time being, therefore, it is convenient to treat all the text in a single Recommendation.

The service requirements for visual telephone services are presented in Recommendation H.200/AV.120-Series; video and audio coding systems and other technical set aspects common to audiovisual services are covered in other Recommendations in the H.200/AV.200-Series.

2. Definitions

BAS: Bit-rate Allocation Signal. Bit position within the frame structure of H.221 to transmit, e.g. commands, control and indication signals, capabilities.

C&I: End-to-end signalling between terminals consisting of "control" which causes a state change in the receiver and "indication" which provides for information as to the functioning of the system, see also H.230.

Data Port: Input/output gate for the user data transmitted within Service Channel or sub-channels according to H.221.

Lip Synchronization: Operation to provide feeling that speaking motion of the displayed person is synchronized with the voice the person makes.

In-band Signalling: Signalling via BAS of the H.221 frame structure.

MCU (Multipoint Control Unit): A piece of equipment located in a node of the network or in a terminal which receives several channels from access ports and, according to certain criterions, processes audiovisual signals and distributes them to the connected channels.

MMI: Man-machine interface between user and terminal/system which consists of a physical section (electro-acoustic, electro-optic transducer, keys, ...) and a logical section dealing with functional operation states.

Narrow-band: Bit rates ranging from 64 kbit/s to 1920 kbit/s. This channel capacity may be provided as a single B/H0/H11/H12 channel or multiple B/H0 channels in ISDN.

Outband Signalling: Signalling via a channel not being part of the B/H0/H11/H12 channel (due to I.400-Series).

Visual Telephone Services: A group of audiovisual services including videophone defined in F.721 and videoconferencing to be defined in H.200/AV.112.

3. System description

3.1 Block diagram and identification of elements

A generic visual telephone system is shown in Figure 1/H.320. It consists of terminal equipment, network, Multipoint Control Unit (MCU) and other system operation entities.

A configuration of the terminal equipment consisting of several functional units is also shown in Figure 1/H.320. "Video 1/0 equipment" includes cameras, monitors and video processing units to provide functions such as split-screen scheme. "Audio 1/0 equipment" includes microphones, loud-speakers and audio processing units to provide such functions as acoustic echo cancellation. "Telematic equipment" are visual aids such as electronic blackboard, still picture transceiver to enhance basic visual telephone communication. "System control" unit carries out such functions as network access through "End-to-network signalling" and end-to-end control to establish common mode of operation and signalling for proper operation of the terminal through "End-to-end signalling". "Video codec" carries out redundancy reduction coding and decoding for video signals, while "audio codec" does the same thing for audio signals. "Delay" in the audio path compensates video codec delay to maintain lip synchronization. "Mux/dmux" unit multiplexes transmitting video, audio, data and control signals into a single bit stream and demultiplexes a received bit stream into consisting multimedia signals. "Network interface" makes necessary adaptation between the network and the terminal according to the user-network interface requirements defined in the I.400-Series.

3.2 Signals

Visual telephone signals are classified into video, audio, data and control as follows:

- audio signals are continuous traffic and require real-time transmission;

<u>Note</u> - In order to reduce the average bit rate of audio signals, voice activation can be introduced (in which case the audio signals are no longer continuous).

- video signals are also continuous traffic, the bit rate allocated to video signals should be as high as possible, in order to maximize the quality within the available channel capacity;

- data signals include still pictures, facsimile and documents, or other facilities, this signal may occur only occasionally as required and may temporarily displace all or part of the audiovisual signal content. It should be noted that data signals are associated only with optional enhancements to the basic visual telephone system, therefore, the opening of a path to carry such signals is preceded by negotiation between the terminals;

- control signals are some system control signals by definition. The path for the terminal-to-network control signals is provided in the D-channel, while the path for the terminal-to-terminal control signals is provided in BAS or Service Channel only when necessary by the mechanism defined in H.221.

3.3 Bit rate options and infrastructure

3.3.1 Communication modes of visual telephone

Communication modes of visual telephone are defined in Table 1 according to their channel configuration and coding.

3.3.2 Terminal types of visual telephone

Table 2 lists terminal types of visual telephone. The terminal type is categorized according to the communication modes and the type of communication channels with which the terminal can communicate; m x B (Type X with parameter a-f), n x H0 (Type Y with parameter 1-5, note), H11/H12 (Type Z with parameter α-ß) or their combinations.

Note - Type Y terminals must have the H0-6B compatibility mode defined in Recommendation H.221 for interworking of evolving networks.

Examples

- Type Xb_3 is a terminal capable of operating at modes a_0, b_1, b_2 and b_3 through B or 2 x B channel.
- Type Xb_3Y1 is a terminal capable of operating at modes a_0, a_1, b_1, b_2, b_3 and g through B, 2 x B or H0 channel.
- Type XfY4Z α is a terminal capable of operating at modes a_0 - k through (1-6) x B, (1-4) x H0 or H11 channel.

For M x B and N x H0 categories, the terminal should be able to operate at all the values of m and n not higher than M and N in principle (note). The type of remote terminal is identified through the transfer rate capability exchange defined in H.242.

Note - Until H.200/AV.254 is recommended, exceptions may arise.

3.3.3 Video codec

As per Recommendation H.261.

3.3.4 Audio codec

As per Recommendations G.711, G.722, H.200/AV.254, AV.253, see Table 1/H.320.

3.3.5 Frame structure

As per Recommendation H.211.

3.3.6 C&I

Identified subset of H.230 is used, see § 4.4.

3.3.7 Communication procedure

As per Recommendation H.242.

3.4 Call control arrangements

To establish intercommunication between various audiovisual terminals, it is necessary to carry out in-band and out-band procedures according to Recommendation H.242 and other relevant CCITT Recommendations.

The different stages of the call are referred according to a point-to-point configuration where terminal X is the calling terminal and Y the called terminal.

3.4.1 Establishment of a visual telephone call - Normal procedure

The provision of the communication is made in the main following steps:

- Phase A: Call set-up, out-band signalling;
- Phase B1: Mode initialization on initial channel;
- Phase CA: Call set-up of additional channel(s), if relevant;
- Phase CB1: Initialization on additional channel(s);
- Phase B2 (or CB2): Establishment of common parameters;
- Phase C: Visual telephone communication;
- Phase D: Termination phase;
- Phase E: Call release.

Phase A - Call set-up

After user initialization, the terminal X performs a call set-up procedure. As soon as the terminal receives an indication from the network that the connection is established, a bidirectional channel is opened from end-to-end, and it overlays H.221 framing on the channel.

Following the connection establishment, all the terminals will start to work in the following modes defined in H.221:

- Type X: Mode OF (A-law or μ-law);
- Type Y and Type Z: Mode OF (A-law or μ-law) audio only.

In-band procedure is activated.

Phase B1 - Mode initialization

Phase B1-1

Using the procedures provided in H.242, framed PCM audio is transmitted in both directions, after frame and multiframe alignment terminal capabilities are exchanged.

Phase B1-2 (terminal procedure)

Determination of the appropriate mode to be transmitted. This will normally be the highest common mode (see Table 3/H.320 for the case using a B or 2 x B channel), but a lower compatible mode could be chosen instead.

In the case that both terminals have announced the capability to work on additional channel(s), terminal X initiates the request for supplementary call set-up. Alternatively, this action may be suspended until the user at X has given the go-ahead, the Y user may also control the additional channel requests. It is for further study.

Note - If the user at either terminal does not wish the call to proceed to two or more channels, even though his terminal has this capability, he must set the terminal such that only single-channel capability is declared in Phase B1-1. In that case, we should distinguish the active capability, wished by the users, from the virtual capability of the terminal.

Phase B1-3 (Mode switching)

Both terminals switch to the mode they have identified in Phase B1-2, using the procedure of H.242.

Note - If the terminals have not both adopted the common mode, an asymmetric communication may result.

Phase CA - Call set-up of the additional channel(s)

Following Phase B1-3 and Phase B2 if relevant, the communication Phase C proceeds on that channel. If additional channels have been requested these are again Phase A (hence the nomenclature "Phase CA"), exactly as in Phase A above, and additional call set-ups are performed by the terminals. On each of the established channels H.221 framing is overlaid (note).

Note - During Phase CA an intermediate audiovisual mode could be offered on the initial channel used for initialization, until full completion of initialization phase.

Phase CB1 - Mode initialization on additional channel(s)

Phase CB1-11

Using the procedure provided in H.242, frame and multiframe alignments are gained.

Phase CB1-12

Synchronization of the channels is achieved.

Phase CB1-2 (Terminal procedure)

Determination of the appropriate mode to be transmitted. This will normally be the highest common mode, but a lower compatible mode could be chosen instead.

Phase CB1-3 (Mode switching)

Both terminals switch to the mode they have identified in Phase B1-2 using the procedure of H.242.

Note - Here again, if the terminals have not both adopted the common mode, an asymmetric communication will result.

Phase B2 (or CB2) - Establishment of common parameters

This phase establishes common operational parameters specific to visual telephone (e.g. encryption) after Phase B1 process is finished. Capabilities or requirements of the receiving side are first indicated then the sending side decides operational parameters and controls the receiving side. BAS codes for this purpose are defined in H.221.

Phase C - Visual telephone communication

In the case where more than one channel is used, there will be intermediate Phases CA, CB1, CB2 as described in this section. Likewise, if additional channels are dropped during the call there will be intermediate Phases CD, CE as described in § 3.4.4. The provisions of this section apply to any channel, initial or additional, for which Phases B1 and B2 have been completed and Phase D not yet started.

Mode switching: According to action by either user (for example, starting a facsimile machine) a different mode from the highest common mode may become more appropriate. Switching to this mode is made according to the procedure of H.242.

Capability change: The user may change the capability of his terminal during the call (for example, by connecting or switching-on auxiliary Telematic equipment); the terminal must initiate the capability exchange procedure defined in H.242.

Phase D - Termination phase

Phase D1 (Terminal procedure)

When one of the users hangs up, the terminal involves Phase D2 directly.

Phase D2 (Mode switching)

Mode OF is forced according to H.242 (or taking into account the result of Phase D1 if different, for further study).

Phase E - Call termination (release)

The terminal which has initiated the hang up sends messages over the D-channel with respect to all channels and idles all of them (that means no more information sent over).

At the other terminal, the first disconnect message received will idle all channels.

The actual disconnection occurs at reception of the other disconnect message(s).

3.4.2 Exceptional procedures for Phases A and B

In case of unsuccessful outcome during Phases A and B (due to many causes), exceptional procedures are provided in order to ensure a suitable service. The matter is for further study.

3.4.3 Exceptional procedures during Phase C

During the actual exchange of audiovisual data, problems may occur in some channels. Fallback procedures, managed by the terminal are activated. The description of the procedures and the appropriate indications are for further study.

3.4.4 Addition and dropping of channels during a visual telephone call

Addition

According to action by a user (for example the activation of auxiliary equipment) one or more additional channels are requested. The procedure follows those described for Phases CA and CB1.

Dropping

Two phases are envisaged:

Phase CD1: The common mode, appropriate to the channel(s) which remains, is selected.

Phase CD2: The mode switching procedure of H.242 is applied to involve the mode identified in Phase CD1; the remaining channel is the channel used for initialization (see Phase A). It supports an appropriate fallback mode. The matter is for further study.

3.4.5 Transmission and display of pictures at the start of a visual telephone call

According to the chosen terminal procedures, pictures may or may not be visible to both users as soon as initialization is complete. In the case that either Phase B1-3 or Phase CB1-3 has activated a common mode, including video, mutual visibility of the users is possible.

The following paragraphs collect alternative procedures which can be used to suspend picture display until user intervention (by mutual agreement or otherwise) and causes pictures to be displayed.

1) No video transmitted: In Phase B1-2 and (if relevant) Phase CB1-2 the mode selected includes "video OFF". During Phase C either user may unilaterally switch to "video ON", alternatively, his terminal may send the C&I BAS code VIR (video indicate ready-to-activate), but not switch to video-ON until VIR is also received from the other terminal. While the incoming video-OFF state remains, the visual telephone screen should display a symbol or message indicating this (i.e. there is no fault).

As already noted in § 3.4.1, Phase B1-2, the request for additional channel may, according to terminal procedure, be delayed while video-OFF is maintained; user action to activate video would then result in procedure Phases CA1, CB1, (CB2 if required).

2) Video pattern transmitted: An electronically generated or other pattern is transmitted instead of the signal from a normal camera. The C&I BAS code VIS (video indicate suppressed) is used to indicate the situation to the remote party.

3) Video transmitted but not displayed: Terminal procedures simply involve local action to display not the incoming signal but an explanatory symbol or message. User action would cause the incoming signal to be displayed, but if this should depend on mutual action by both users then a new C&I BAS code VRD (video ready-to-display) must be defined. This point is for further study.

3.5 Optional enhancements

3.5.1 Data ports

Data ports as physical I/O ports of the terminal for Telematic and other equipment are activated/deactived by BAS commands. Depending on the transmission capability of a connection, e.g. multiples of B/H0 channels etc., various bit rates are available at these ports. Allocation of bit streams to the port(s) is performed by in-band signalling. Data conveyed at the port(s) is transparent, data rates being listed in H.221, Annex 1.

3.5.2 Encryption

Encryption may be applied on audio and video signals separately (preferably for multipoint connections) or on audio and video signals multiplexed. Switching-on and off the encryption process has to be signalled between the terminals (or terminal and MCU respectively) via in-band signalling.

4. Terminal requirements

4.1 Environments

Under study.

4.2 Audio and video arrangements

Under study.

4.3 Delay compensation in the audio path

The H.261 video codecs require some processing delay, while the H.200/AV.250-Series audio codecs involve much less delay. Hence, if lip synchronization is to be maintained, that video processing delay must be compensated in the audio path. Since video coder and decoder delays may vary according to implementation, delay compensation must be carried out individually at the coder and decoder. A reference measurement method of video coder and decoder delays is defined in Recommendation H.261.

4.4 Control and indications (C&I)

C&I are chosen from the general audiovisual set contained in Recommendation H.230. For visual telephone systems, those signals in Table 4/H.320 are used, where their source, sink, synchronization with picture, transmission channel and codewords are indicated.

All visual telephone terminals have a video source providing a picture of participants, and some terminals may have additional video sources; the participant-picture source is designated #1, having the associated symbol VIA. When incoming video is "ON" (BAS command (010) [1 or 2]) and VIA, VIA2, VIA3 have not been transmitted, source #1 is assumed.

5. Intercommunications

The mechanisms for intercommunication with other services are described in the H.200/AV.240-Series.

5.1 Intercommunication between different visual telephone terminal types

A common mode of operation is determined as described in § 3.4.1 above. D-channel signalling should include new LLC and HLC which are appropriate for audiovisual services, but this point is for further study.

5.2 Intercommunication with telephony

Note - Description of this section is for communications using a B-channel.

5.2.1 Intercommunication with ISDN telephones

A call from a visual telephone to an ISDN telephone is first placed as an audiovisual call, but the ISDN telephone returns "Incompatible destination" or the network returns "Recovery on timer expiry" in case of no responses from the called side, then the visual telephone may switch to a speech or 7 kHz audio bearer service call.

A call from ISDN telephone to a visual telephone is accepted by the visual telephone because every audiovisual terminal is equipped with this telephone capability as a minimum function.

For both of the above cases, the operational mode of communication is G.711 speech or G.722 audio.

5.2.2 Intercommunication with PSTN telephones

A call from visual telephone to a PSTN telephone may be initiated as an audiovisual call, but the network returns "No route to destination", then the visual telephone may switch to a speech or 3.1 kHz audio bearer service call. The operational mode of communication is G.711 audio coding. A call from a PSTN telephone is routed into the ISDN as a 3.1 kHz audio call which can be responded by the visual telephone for the same reason as described in § 5.2.1. The operational mode of communication is 3.1 kHz audio.

5.3 Intercommunication with other audiovisual terminals

A common mode of operation is determined according to H.242.

6. Maintenance

Some loop-back functions are envisaged to allow verification of the functional aspects of the terminal in order to ensure correct operation of the system and satisfactory quality of the service to the remote party. The following loop-back functions (see Figure 2/H.320) are envisaged:

- Loop at terminal-network interface (towards network)

  Upon receiving the "digital loop back" BAS, loop back is activated at the digital interface of the terminal toward the network side. In case of a multiple B/HO channel arrangement, loop back is activated in each connection.

- Loop at terminal-network interface (towards terminal)

  The procedure is for further study.

- Loop at analogue I/O interface

  Upon receiving the "video loop back" or "audio loop back" BAS, loop back is activated at the analogue interface of the video/audio codec towards the video/audio codec.

The opportunity of having a self-checking procedure at terminal stage is for further study.

7. Human factor aspects

To achieve error free and uncomplicated utilization of terminal equipment and service from the users standpoint, human factor related aspects have to be studied and recommended. These aspects deal with the flow of information between user and terminal/network. This information can be divided into a physical section and a logical section of the MMI.

Physical section

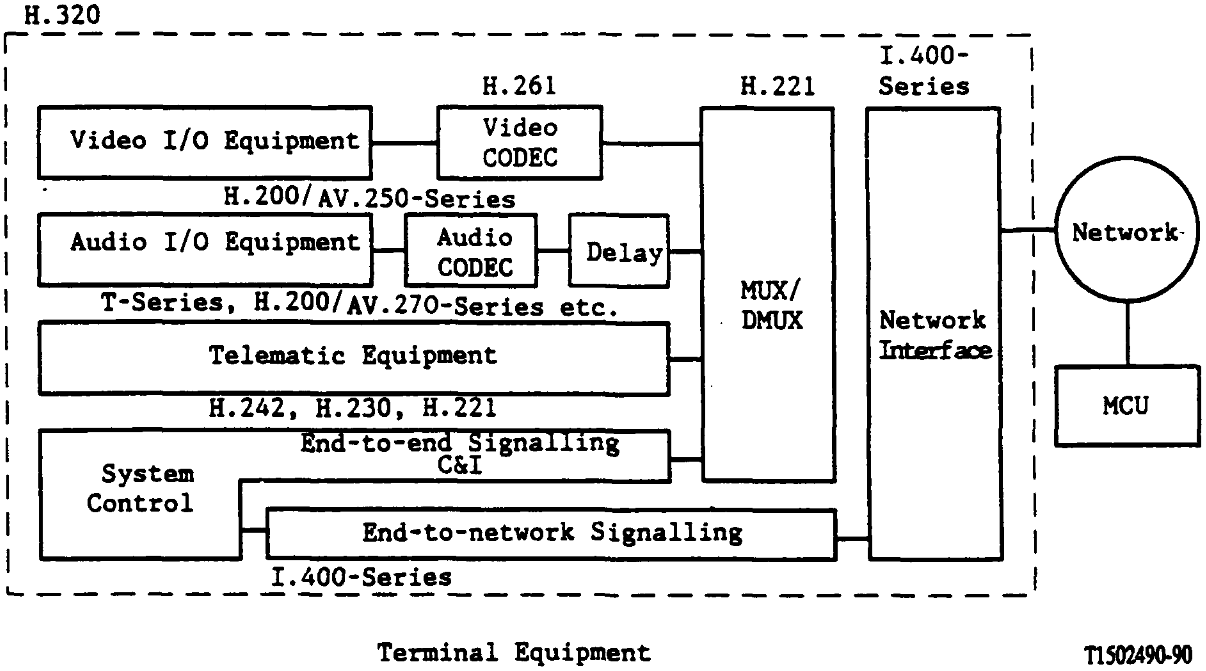
- Figures and properties of transducers (camera, microphone, etc.).

- Signals particularly related to the service, keys, pictograms.

Logical section

- Procedures, e.g. for call establishment/release, during communication phase.

- Consistency between the MMIs of visual telephone and terminals of other teleservices.



MCU: Multipoint Control Unit

FIGURE 1/H.320

Visual telephone system

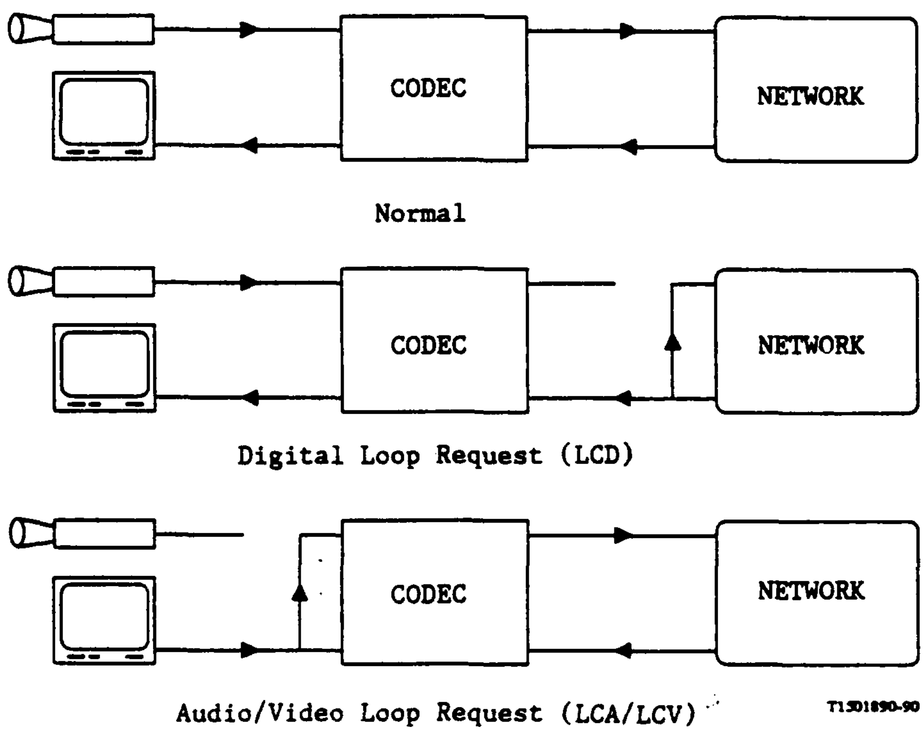


FIGURE 2/H.320

Loop back

TABLE 1/H.320

Communication modes of visual telephone

<table>
<tr><th colspan="2" rowspan="2">Visual Telephone Mode (Note 1)</th><th rowspan="2">Channel Rate (kbit/s)</th><th rowspan="2">ISDN Channel (Note 2)</th><th colspan="2">ISDN Interface</th><th colspan="2">Coding</th></tr>
<tr><th>Basic</th><th>Primary Rate</th><th>Audio</th><th>Video</th></tr>
<tr><td rowspan="2">a</td><td>a0</td><td rowspan="2">64</td><td rowspan="2">B</td><td rowspan="5"></td><td rowspan="16">applicable</td><td>G.711</td><td>not applicable</td></tr>
<tr><td>a1</td><td>H.200/AV.254</td><td rowspan="15">H.261</td></tr>
<tr><td rowspan="3">b</td><td>b1</td><td rowspan="3">128</td><td rowspan="3">2B</td><td>G.711</td></tr>
<tr><td>b2</td><td>G.722</td></tr>
<tr><td>b3</td><td>H.200/AV.254 /253 (Note 1)</td></tr>
<tr><td colspan="2">c</td><td>198</td><td>3B</td><td rowspan="11">not applicable</td><td rowspan="11">G.722</td></tr>
<tr><td colspan="2">d</td><td>256</td><td>4B</td></tr>
<tr><td colspan="2">e</td><td>320</td><td>5B</td></tr>
<tr><td colspan="2">f</td><td>384</td><td>6B</td></tr>
<tr><td colspan="2">g</td><td>384</td><td>H0</td></tr>
<tr><td colspan="2">h</td><td>768</td><td>2H0</td></tr>
<tr><td colspan="2">i</td><td>1152</td><td>3H0</td></tr>
<tr><td colspan="2">j</td><td>1536</td><td>3H0</td></tr>
<tr><td colspan="2">k</td><td>1536</td><td>H11</td></tr>
<tr><td colspan="2">l</td><td>1920</td><td>5H0</td></tr>
<tr><td colspan="2">m</td><td>1920</td><td>H12</td></tr>
</table>

Note 1 - (Audio coding of mode b3) In addition to H.200/AV.254, higher quality audio coding such as H.200/AV.253 may be used for this mode.

Note 2 - For multiple channels of B/H0, all channels are synchronized at the terminal according to § 2.7/H.221.

TABLE 2/H.320

Visual telephone terminal type

Mode	Type X (Note 2)										Type Y (Note 3)					Type Z	
	a	b1	b2	b3	b4	b5	c	d	e	f	1	2	3	4	5	α	β
a0 B(audio only)	X	X	X	X	X	X	X	X	X	X							
a1 B(H.200/AV.254 audio)	X	X	X	X			X	X	X	X							
b1 2B(G.711 audio)		X	X	X	X	X	X	X	X	X							
b2 2B(G.722 audio)			X	X		X	X	X	X	X							
b3 2B(H.200/AV.254 audio)				X			X	X	X	X							
c 3B							X	X	X	X							
d 4B								X	X	X							
e 5B									X	X							
f 6B										X							
g H0											X	X	X	X	X		
h 2H0												X	X	X	X		
i 3H0													X	X	X		
j 4H0														X	X		
k H11																X	
l 5H0															X		
m H12																	X

Note 1 - "X" means the mode is equipped with the terminal of the type.

Note 2 - Types Xb4 and Xb5 are defined to take into account that H.200/AV.254 has not yet been recommended.

Note 3 - Terminal of this type must have the H0-6B compatible mode defined in Recommendation H.221.

APPENDIX B

DRAFT FEDERAL STANDARD 1080

TELECOMMUNICATIONS: VIDEO CODER/DECODER FOR AUDIOVISUAL SERVICES AT 56 TO 1,920 KBIT/S

PROPOSED FEDERAL STANDARD 1080

TELECOMMUNICATIONS: VIDEO CODER/DECODER FOR AUDIOVISUAL SERVICES AT 56 TO 1,920 KBIT/S

Draft 2 - November 8, 1990

This proposed Federal Standard has not yet been approved and is subject to modification

Prepared By:
National Communications System
Office of Technology & Standards

Published By:
General Services Administration
Office of Information Resources Management

NOVEMBER 8, 1990
DRAFT

FOREWORD

This standard is issued by the General Services Administration pursuant to the Federal Property and Administrative Services Act of 1949, as amended.

This document provides Federal departments and agencies a comprehensive description of the performance and interoperability criteria for video coders and decoders (codecs) used in video teleconferencing and video phone applications. This standard was developed within the Federal Telecommunication Standards Committee (FTSC). Standard development was based on the requirements contained in the Statement of Requirements (SOR) for the development of a standard for digital transmission of video teleconferencing.

This standard shall be used by all Federal departments and agencies in the design and procurement of codecs used in video teleconferencing and video phone applications. Neither this nor any other standard in high technology field such as telecommunications can be considered complete and ageless. Periodic revisions will be made as required.

This standard is technically equivalent to CCITT Recommendation H.261 (1990) and ANSI T1.p64x (199x). The following areas have been changed or added to H.261 to meet the needs of the Federal Government: Sections 1, 2, 3, 4, 5, 7.2.2, 9.4.1. In addition to these section changes, editorial changes have been made to reflect common english (e.g. change colour to color).

Recommendation H.261 specifies service from 64 kbit/s through 1,920 kbit/s, and ANSI standard T1.p64-199x specifies service from 56 kbit/s through 1,536 kbit/s. To avoid confusion on applications within the Federal Government involving both national and international interoperability, this standard encompasses both ranges of data rates to specify service from 56 kbit/s through 1,920 kbit/s. It should be noted that most standard data networks in the United States carry data from 56 kbit/s to 1,536 kbit/s.

The digitally encoded video specified by these standard is intended to be used with CCITT Recommendations H.221, Frame Structure for a 64 to 1920 kbit/s channel in Audiovisual Teleservices; H.242, System for Establishing Communication Between Audiovisual Terminals using Digital Channels up to 2 Mbit/s; and H.230, Frame-Synchronous Control and Indication Signals for Audiovisual Systems. It is intented to adopt these recommendations into Federal Standards. Titles and application for Federal use is yet to be determined.

Contents

TELECOMMUNICATIONS: VIDEO CODER/DECODER FOR AUDIOVISUAL SERVICES AT 56 TO 1,920 KBIT/S

1. Scope, Purpose, and Application

1.1 Scope. This standard describes the video coding and decoding methods for the moving picture component of audiovisual service at the rates of 56 to 1,920 kbit/s. Included are specifications for the video source format, the source coding algorithm, the video multiplex arrangement, and the forward error correction code. This standard is technically equivalent to CCITT Recommendation H.261 (1990).

1.2 Purpose. The purpose of this document is to improve the Federal acquisition process by providing Federal departments and agencies a comprehensive, authoritative source for video coders and decoders (codecs) used in video teleconferencing and video phone applications. The technical parameters of this document may be exceeded in order to satisfy certain specific requirements, provided that interoperability is maintained. That is, the capability to incorporate features such as additional standard and nonstandard interfaces is not precluded.

1.3 Application. This standard is intended to assure interoperability among Federal video teleconferencing and video phone systems employing video codecs at rates between 56 kbit/s and 1,920 kbit/s.

Two categories of applications are defined for Federal use. The specifications defined below are a minimum for each category.

Category 1 is for service that is limited to lower data rates, such as provided by a basic rate Integrated Services Digital Network (ISDN). Examples of uses are video telephone or surveillance type services. Codecs acquired for Category 1 applications will be capable of operation at 56 kbit/s and 64 kbit/s (p=1), and optionally at 112 kbit/s and 128 kbit/s (p=2), using Quarter Common Intermediate Format (QCIF).

Category 2 is for service requiring higher quality than Category 1 such as needed for full video teleconferencing. Codecs acquired for Category 2 applications will contain the functionality of Category 1 codecs, plus at a minimum, be capable of operation at 128, 192, 256, and 384 kbit/s (p=2,3,4,5,6). All category 2 codecs must, at a minimum, be capable of operation at Full Common Intermediate Format (CIF) at rates equal to or above 128 kbit/s.

1.3.1 The digitally encoded video is intended to be transmitted within the frame structure described in CCITT Recommendation H.221 for narrowband audiovisual services. This frame structure multiplexes subchannels for audio, video, data, and telematic transmission, as well as in-channel terminal-to-terminal signalling information, within an overall transmission channel of 56 to 1,920 kbit/s.

In an ISDN, the overall transmission channel may consist of 1 to 6 B (64 kbit/s) channels, 1 to 4 H0 (384 kbit/s) channels, an H10 (1,472 kbit/s) channel, or an H11 (1,536 kbit/s) channel. The framed video signal can also be carried on other switched or dedicated digital transmission facilities, such as 1 to 6 56 kbit/s connections, a DS1 connection, or a fractional DS1 connection. The H.221 frame structure provides for the synchronization of multiple connections.

NOTE: The video coding algorithm described in this standard is a variable-rate algorithm. Video transmission is not fixed at multiples of 56 or 64 kbit/s, but instead occupies whatever bandwidth allocated for video within an overall audiovisual communications system. "Px64 kbit/s" are the nominal transmission rates of the overall system. CCITT Recommendation H.221 provides for operating at multiples of 56 and 64 kbit/s.

CCITT Recommendation H.242 describes the in-channel terminal-to-terminal communication control procedures. These procedures allow audiovisual terminals with different capabilities to interwork with each other and with existing telephone equipment. These procedures also allow terminals to switch among compatible modes of operation to support additional applications, for example, sending a facsimile or connecting two personal computers.

Additional frame-synchronous control and indication signals such as freeze picture, video loopback, and simple multipoint controls are described in CCITT Recommendation H.230.

1.3.2 This standard shall be used by all Federal departments and agencies in the design and procurement of video teleconferencing and video phone systems employing video codecs. This standard is mandatory only for those video codecs operating at rates between 56 kbit/s and 1,920 kbit/s. The standard shall be used in the planning, design, and procurement, including lease and purchase, of all new video communications systems that utilize video codecs. It is not intended that existing equipment and systems be immediately converted to comply with the provisions of this standard. New equipment and systems and those undergoing major modification or rehabilitation shall conform to this standard.

1.3.3 For application of this standard within the Department of Defense, users should refer to the supplemental requirements contained in Military Standard 188-131.

2. Effective Date. The use of this standard by U.S. government departments and agencies is mandatory, effective 180 days following the date of this standard.

3. Changes. When a Government department or agency considers that this standard does not provide for its essential needs, a statement citing specific requirements shall be sent in duplicate to the General Services Administration (K), Washington, DC 20405, in accordance with the provisions of Federal Property Management Regulation 41 CFR 101-29.403.1. The General Services Administration will determine the appropriate action to be taken and will notify the agency.

PREPARING ACTIVITY:

National Communications System
Office of Technology and Standards
Washington, DC 20305-2010

4. Referenced and Related Standards

4.1 Referenced Standards. This standard is intended for use in conjunction with the following standards:

CCIR Recommendation 601-1, Encoding Parameters of digital Television for Studios, Recommendations and Reports of the CCIR, Volume XI, Part 1, 1986

CCITT Recommendation H.221, Frame Structure for a 64 to 1,920 kbit/s Channel in Audiovisual Teleservices, 1990

CCITT Recommendation H.230, Frame-Synchronous Control and Indication Signals for Audiovisual Systems, 1990

CCITT Recommendation H.242, System for Establishing Communication Between Audiovisual Terminals Using Digital Channels up to 2 Mbit/s, 1990

4.2 Related Standards. The standards listed here are for information only and are not essential for the completion of the requirements of this standard.

ANSI T1.306-1990, American National Standard for Telecommunications - Digital Processing of Audio Signals - Algorithm and Line Format for Transmission of 7-kHz Audio Signals at 64/56 kbit/s

ANSI T1.p64-199x, American National Standard for Telecommunications - Video Coder/Decoder for Audiovisual Services at 56 to 1,536 kbit/s

CCITT Recommendation H.261, Video Codec for Audiovisual Services at px64 kbit/s, 1990

CCITT Recommendation H.320, Narrowband Visual Telephone Systems and Terminal Equipment, 1990

MIL-STD-188-131, Video Codec Equipment for Video Teleconferencing Applications, 1990

5. Abbreviations

ANSI	American National Standards Institute
BCH	Bose-Chaudhuri-Hocquenghem
CBP	Coded-Block Pattern
CCIR	International Radio Consultative Committee
CCITT	International Telegraph and Telephone Consultative Committee
CIF	Common Intermediate Format
CODEC	Coder/Decoder
DCT	Discrete Cosine Transform
EOB	End of Block
FIL	Loop Filter
FLC	Fixed Length Code
GBSC	Group of Blocks Start Code
GEI	GOB Extra Insertion information
GN	Group Number
GOB	Group of Blocks
GQUANT	GOB Quantizer information
GSPARE	GOB Spare information
HRD	Hypothetical Reference Decoder
IDCT	Inverse Discrete Cosine Transform
INTER	Inter-picture prediction
INTRA	Intra-picture prediction
MB	Macroblock
MBA	Macroblock Address
MC	Motion Compensation
MQUANT	Macroblock Quantizer
MTYPE	Macroblock Type information
MVD	Motion Vector Data
PEI	Picture Extra Insertion information
PSC	Picture Start Code
PSPARE	Picture Spare information
PTYPE	Picture Type information
QCIF	Quarter-CIF
QUANT	Quantizer
REC	Reconstruction level
TCOEFF	Transform Coefficient
TR	Temporal Reference

VLC Variable Length Code

6. Brief Specification

An outline block diagram of the codec is given in Figure 1.

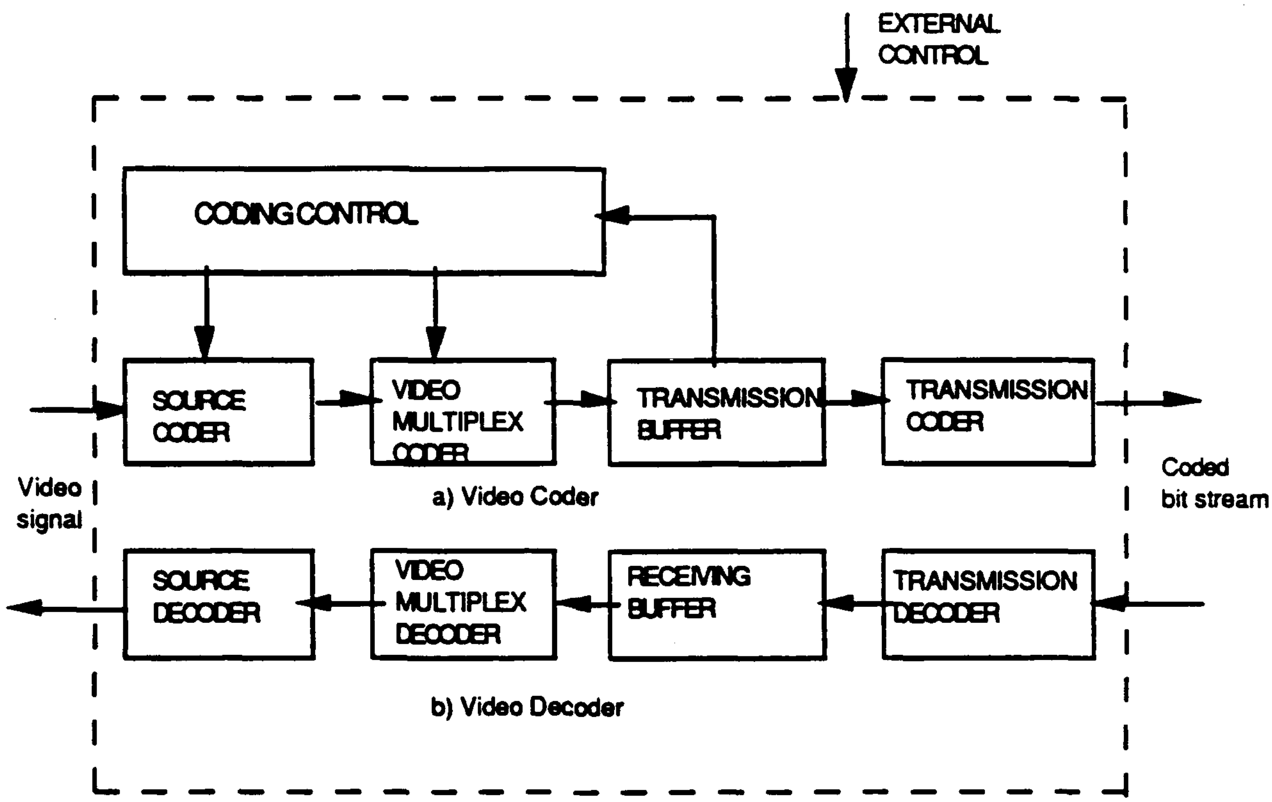


Figure 1
Outline block diagram of the video codec

6.1 Video Input and Output. To permit a single recommendation to cover use in and between regions using 625 and 525 line television standards, the source coder operates on pictures based on a common intermediate format (CIF). The standards of the input and output television signals, which may, for example, be composite or component, analogue or digital and the methods of performing any necessary conversion to and from the source coding format are not subject to recommendation.

6.2 Digital Output and Input. The video coder provides a self-contained digital bit stream which may be combined with other multi-facility signals (for example as defined in Rec. H.221). The video decoder performs the reverse process.

6.3 Sampling Frequency. Pictures are sampled at an integer multiple of the video line rate. This sampling clock and the digital network clock are asynchronous.

6.4 Source Coding Algorithm. A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The decoder has motion compensation capability, allowing

6.5 Bit Rate. This Recommendation is primarily intended for use at video bit rates between approximately 40 kbit/s and 2 Mbit/s.

6.6 Symmetry of Transmission. The codec may be used for bidirectional or unidirectional visual communication.

6.7 Error Handling. The transmitted bit-stream contains a BCH(511,493) Forward Error Correction code. Use of this by the decoder is optional.

6.8 Multipoint Operation. Features necessary to support switched multipoint operation are included.

7. Source Coder

7.1 Source Format. The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is +/-50 ppm.

Pictures are coded as luminance and two color difference components (Y, C_B and C_R). These components and the codes representing their sampled values are defined in CCIR Recommendation 601.

Black = 16
White = 235
Zero color difference = 128
Peak color difference = 16 and 240

These values are nominal ones and the coding algorithm functions with input values of 1 through to 254. See table 6.

Two picture scanning formats are specified.

In the first format (CIF), the luminance sampling structure is 352 pels per line, 288 lines per picture in an orthogonal arrangement. Sampling of each of the two color difference components is at 176 pels per line, 144 lines per picture, orthogonal. Color difference samples are sited such that their block boundaries coincide with luminance block boundaries as shown in Figure 2. The picture area covered by these numbers of pels and lines has an aspect ratio of 4:3 and corresponds to the active portion of the local standard video input.

Note: The number of pels per line is compatible with sampling the active portions of the luminance and color difference signals from 525 or 625 line sources at 6.75 and 3.375 MHz respectively. These frequencies have a simple relationship to those in CCIR Recommendation 601.

The second format, Quarter-CIF (QCIF), has half the number of pels and half the number of lines stated above. All codecs must be able to operate using QCIF. Some codecs can also operate with CIF.

Means shall be provided to restrict the maximum picture rate of encoders by having at least 0, 1, 2, or 3 non-transmitted pictures between transmitted ones. Selection of this minimum number and CIF or QCIF shall be by external means (for example via Recommendation H.221).

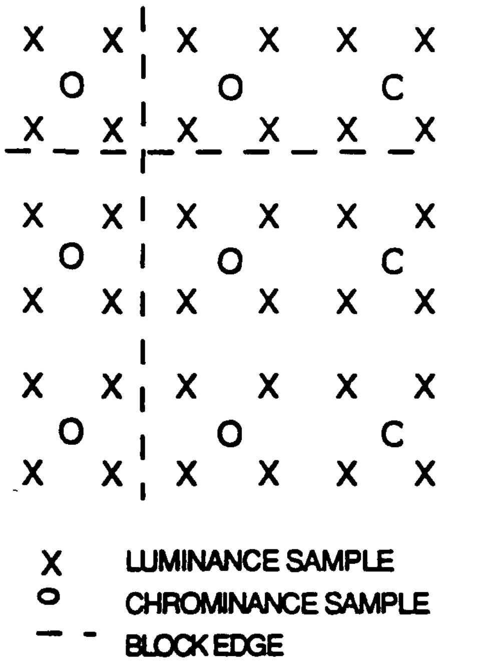
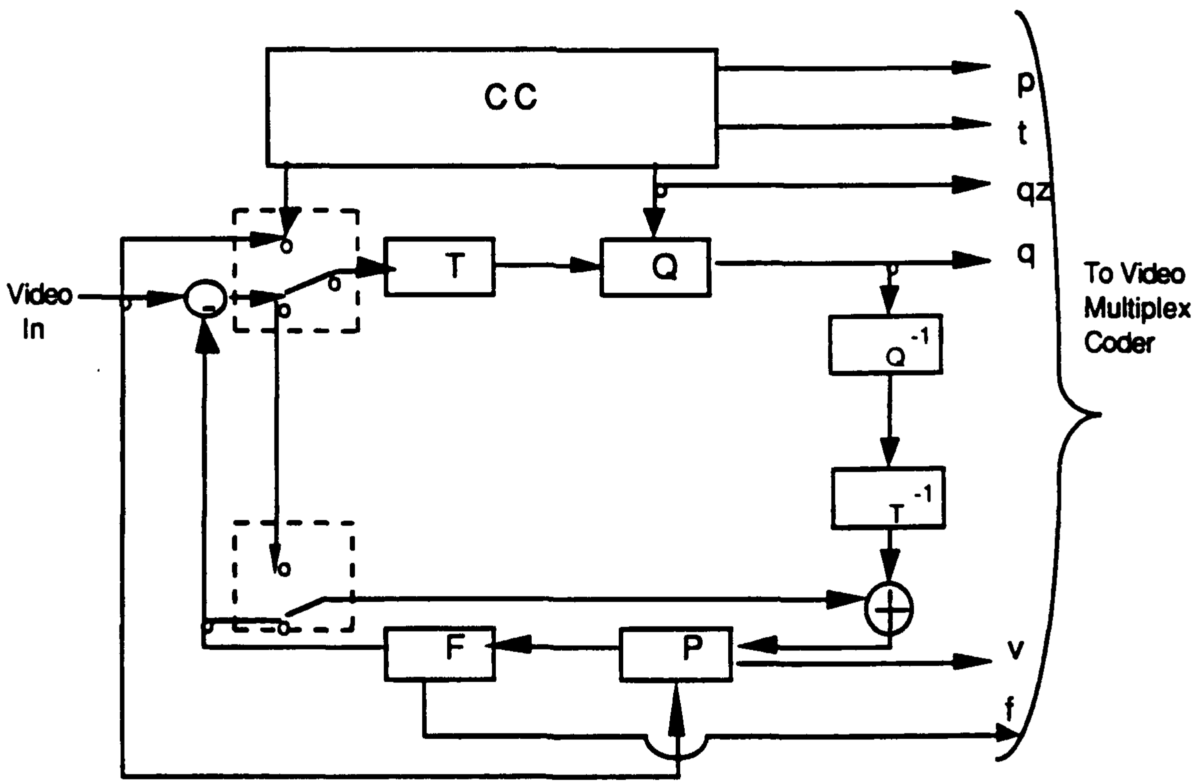


Figure 2
Positioning of luminance and chrominance samples

7.2 Video Source Coding Algorithm. The source coder is shown in generalized form in Figure 3. The main elements are prediction, block transformation, and quantization.

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 pel line blocks which are segmented as transmitted or non-transmitted. Further, four luminance blocks and the two spatially corresponding color difference blocks are combined to form a macroblock as shown in Figure 10 of 8.2.4.

The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the coding control strategy. Transmitted blocks are transformed and resulting coefficients are quantized and variable length coded.



T Transform
Q Quantizer
P Picture Memory with motion compensation variable delay
F Loop filter
CC Coding control

p Flag for INTRA/INTER
t Flag for transmitted or not
qz Quantizer indication
q Quantizing index for transform coefficients
v Motion vector
f Switching on/off of the loop filter

Figure 3
Source coder

7.2.1 Prediction. The prediction is inter-picture and may be augmented by motion compensation (see 7.2.2) and a spatial filter (see 7.2.3).

7.2.2 Motion Compensation. Motion Compensation (MC) is optional in the encoder, but will be considered in evaluating the price-performance characteristics of competing vendor proposals. The decoder will accept one vector per macroblock. Both horizontal and vertical components of these motion vectors have integer values not exceeding +/-15. The vector is used for all four luminance blocks in the macroblock. The motion vector for both color difference blocks is derived by halving the component values of the macroblock vector and truncating the magnitude parts toward zero to yield integer components.

A positive value of the horizontal or vertical component of the motion vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right or below the pels being predicted.

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

7.2.3 Loop Filter. The prediction process may be modified by a two-dimensional spatial filter (FIL) which operates on pels within a predicted 8 by 8 block. The filter is separable into one dimensional horizontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4 except at block edges where one of the taps would fall outside the block. In such cases the 1-D filter is changed to have coefficients of 0, 1, 0. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is greater or equal to one half are rounded up.

The filter is switched on/off for all 6 blocks in a macroblock according to the macroblock type. (See 8.2.3 MTYPE).

7.2.4 Transformer. Transmitted blocks are first processed by a separable 2-dimensional Discrete Cosine Transform of size 8 by 8. The output from the inverse transform ranges from -256 to +255 after clipping to be represented with 9 bits. The transfer function of the inverse transform is given by:

$$f(x,y) = 1/4 \sum_{u=0}^{7} \sum_{v=0}^{7} C(u)\, C(v)\, F(u,v) \cos[Pi(2x+1)u/16] \cos[Pi(2y+1)v/16]$$

with u, v, x, y = 0, 1, 2, ..., 7

where x,y = spacial coordinates in the pel domain
u,v = coordinates in the transform domain
C(u) = 1/SQRT(2) for u = 0, otherwise 1.
C(v) = 1/SQRT(2) for v = 0, otherwise 1.

Note: Within the block being transformed, x=0 and y = 0 refer to the pel nearest the left and top edges of the picture respectively.

The arithmetic procedures for computing the transforms are not defined, but the inverse transform shall meet the error tolerance specified in section 10.

7.2.5 Quantization. The number of quantizers is 1 for the INTRA DC coefficient and 31 for all other coefficients. Within a macroblock the same quantizer is used for all coefficients except the INTRA DC one. The decision levels are not defined. The INTRA DC coefficient is nominally the transform value linearly quantized with a stepsize of 8 and no dead-zone. Each of the other 31 quantizers is also nominally linear but with a central dead-zone around zero and with a stepsize of an even value in the range 2 to 62.

The reconstruction levels are as defined in 8.2.4.

Note: For the smaller quantization step sizes, the full dynamic range of the transform coefficients cannot be represented.

7.2.6 Clipping of Reconstructed Picture. To prevent quantization distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. The clipping function is applied to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

7.3 Coding Control. Several parameters may be varied to control the rate of generation of coded video data. These include processing prior to the source coder, the quantizer, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures.

7.4 Forced Updating. This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update pattern is not defined. For control of accumulation of inverse transform mismatch error a macroblock should be forcibly updated at least once per every 132 times it is transmitted.

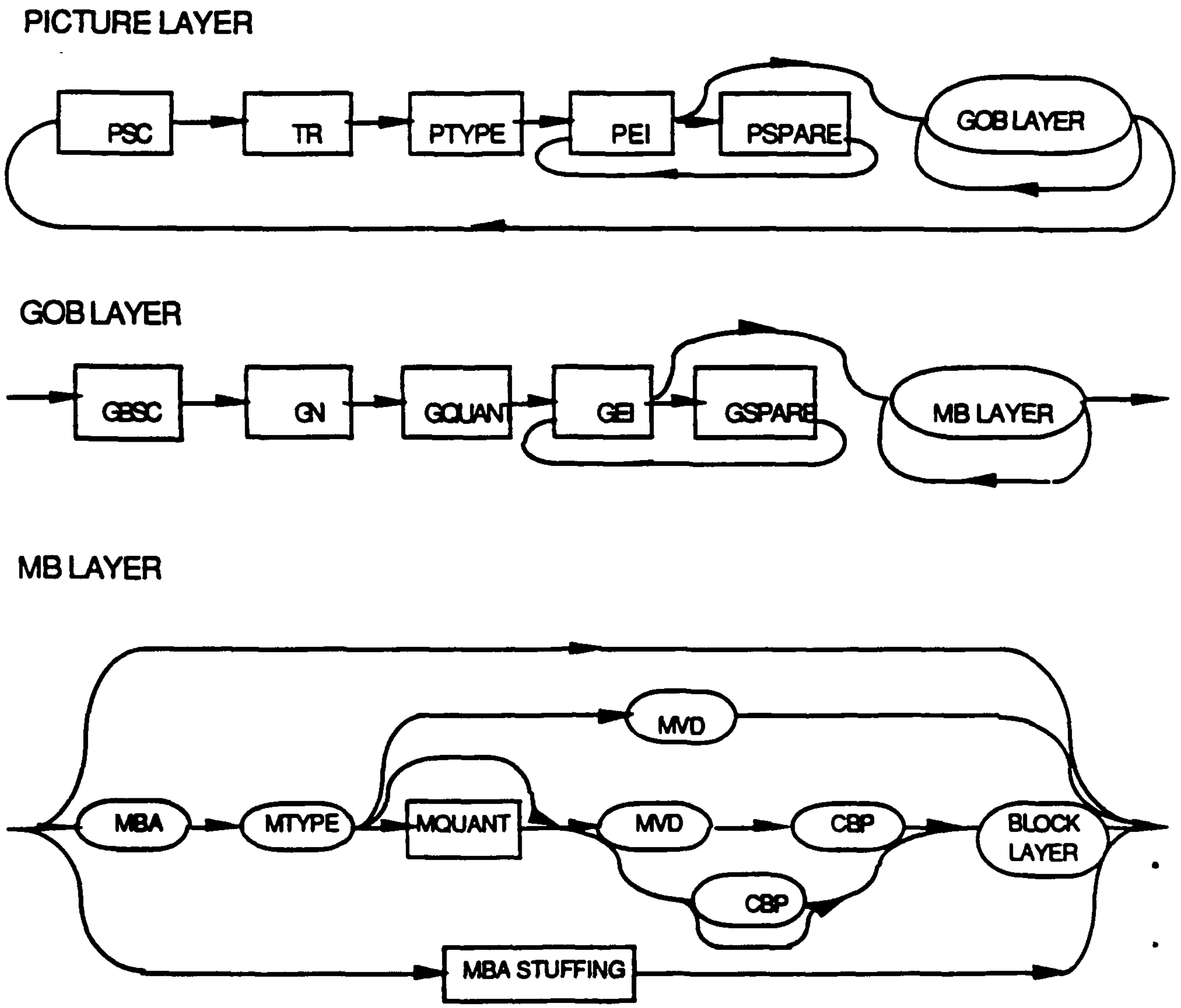
8. Video Multiplex Coder

8.1 Data Structure. Unless specified otherwise the most significant bit is transmitted first. This is Bit 1 and is the leftmost bit in the code tables in this document. Unless specified otherwise all unused or spare bits are set to '1'. Spare bits must not be used until their functions are specified by CCITT.

8.2 Video Multiplex Arrangement. The video multiplex is arranged in a hierarchical structure with four layers. From top to bottom the layers are:

Picture
Group of Blocks (GOB)
Macroblock (MB)
Block

A syntax diagram of the video multiplex coder is shown in Figure 4. Abbreviations are defined in section 5 and in later sections.



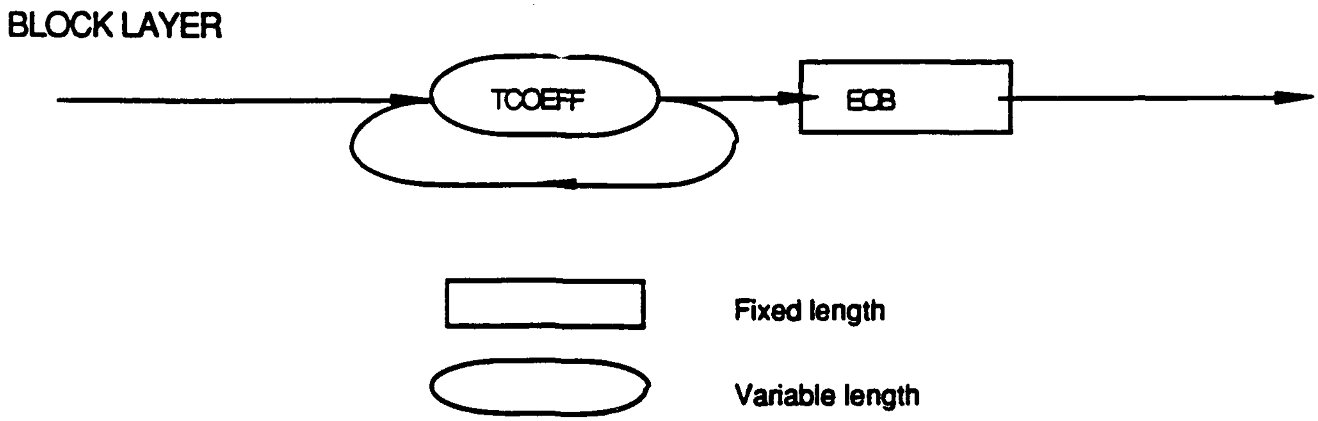


Figure 4
Syntax diagram for the video multiplex coder

8.2.1 Picture Layer. Data for each Picture consists of a Picture Header followed by data for GOBs. The structure is shown in Figure 5. Picture Headers for dropped pictures are not transmitted.

Picture Start Code (PSC): 20 bits

A word of 20 bits. Its value is 0000 0000 0000 0001 0000

Temporal Reference (TR): 5 bits

A five bit number which can have 32 possible values. It is formed by incrementing its value in the previously transmitted Picture Header by 1 plus the number of non-transmitted pictures (at 29.97 Hz) since that last transmitted one. The arithmetic is performed with only the 5 LSBs.

Type Information (PTYPE): 6 bits

Information about the complete picture;

Bit 1 Split screen indicator.
 '0' off. '1' on.
Bit 2 Document camera indicator.
 '0' off, '1' on.
Bit 3 Freeze Picture Release.
 '0' off, '1' on.
Bit 4 Source Format.
 '0' QCIF, '1' CIF.
Bits 5 to 6 Spare.

Extra Insertion Information (PEI): 1 bit

A bit which when set to '1' signals the presence of the following optional data field.

Spare Information (PSPARE): 0/8/16...bits

If PEI is set to '1', then 9 bits follow consisting of 8 bits of data (PSPARE) and then another PEI bit to indicate if a further 9 bits follow and so on. Encoders must not insert PSPARE until specified by CCITT. Decoders must be designed to discard PSPARE if PEI is set to 1. This will allow CCITT to specify future "backward" compatible additions in PSPARE.

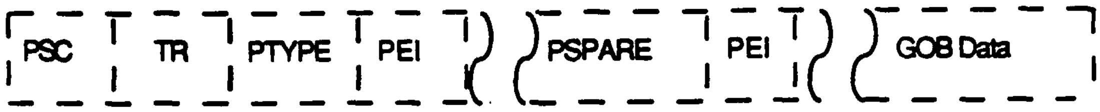


Figure 5
Structure of Picture layer

8.2.2 Group of Blocks Layer. Each picture is divided into Groups of Blocks (GOBs). A group of blocks (GOB) comprises one twelfth of the CIF or one third of the QCIF picture areas (see Figure 6). A GOB relates to 176 pels by 48 lines of Y and the spacially corresponding 88 pels by 24 lines of each of C_B and C_R.

Data for each Group of Blocks consists of a GOB Header followed by data for macroblocks. The structure is shown in Figure 7. Each GOB Header is transmitted once between Picture Start Codes in the CIF or QCIF sequence numbered in Figure 6, even if no macroblock data is present in that GOB.

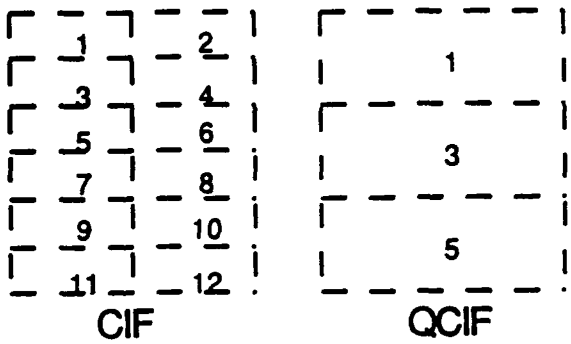


Figure 6
Arrangement of GOBs in a Picture

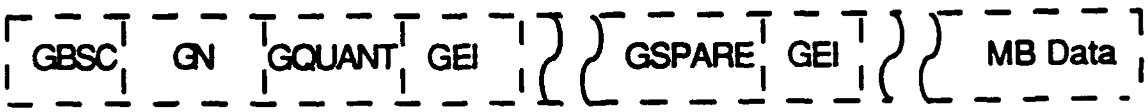


Figure 7
Structure of Group of Blocks Layer

Group of Blocks Start Code (GBSC): 16 bits

A word of 16 bits, 0000 0000 0000 0001.

Group Number (GN): 4 bits

Four bits indicating the position of the groups of blocks. The bits are the binary representation of the numbers in Figure 6. Group numbers 13, 14 and 15 are reserved for future use. Group number 0 is used in the PSC.

Quantizer Information (GQUANT): 5 bits

A fixed length of codeword of 5 bits which indicates the quantizer to be used in the group of blocks until overridden by any subsequent MQUANT. The codewords are the natural binary representations of the values of QUANT (see 8.2.4.) which, being half the stepsizes, range from 1 to 31.

Extra Insertion Information (GEI): 1 bit

A bit which when set to '1' signals the presence of the following optional data field.

Spare Information (GSPARE): 0/8/16... bits

If GEI is set to '1', then 9 bits follow consisting of 8 bits of data and then another GEI bit to indicate if a further 9 bits follow and so on. Encoders must not insert GSPARE until specified by CCITT. Decoders must be designed to discard GSPARE if GEI is set to 1. This will allow CCITT to specify future "backward" compatible additions in GSPARE.

Note: Emulation of start codes may occur if the future specification of GSPARE has no restrictions on the final GSPARE data bits.

8.2.3 Macroblock Layer. Each GOB is divided into 33 macroblocks as shown in Figure 8. A macroblock relates to 16 pels by 16 lines of Y and the spatially corresponding 8 pels by 8 lines of each of C_B and C_R.

1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

Figure 8
Arrangement of Macroblocks in GOB

Data for a macroblock consists of a MB Header followed by data for blocks (Figure 9). MQUANT, MVD and CBP are present when indicated by MTYPE.

MBA	MTYPE	MQUANT	MVD	CBP	Block Data

Figure 9
Structure of Macroblock layer

Macroblock Address (MBA)
Variable length

A variable length codeword indicats the position of a macroblock within a group of blocks. The transmission order is shown in Figure 8. For the first transmitted macroblock in a GOB, MBA is the absolute address in Figure 8. For subsequent macroblocks, MBA is the difference between the absolute addresses of the macroblock and the last transmitted macroblock. The code table for MBA is given in Table 1.

An extra codeword is available in the table for bit stuffing immediately after a GOB header or a coded macroblock (MBA Stuffing). This codeword should be discarded by decoders.

The VLC for start code is also shown in Table 1.

MBA is always included in transmitted macroblocks.

Macroblocks are not transmitted when they contain no information for that part of the picture.

Table 1
VLC Table for Macroblock Addressing

MBA	CODE	MBA	CODE
1	1	17	0000 0101 10
2	011	18	0000 0101 01
3	010	19	0000 0101 00
4	0011	20	0000 0100 11
5	0010	21	0000 0100 10
6	0001 1	22	0000 0100 011
7	0001 0	23	0000 0100 010
8	0000 111	24	0000 0100 001
9	0000 110	25	0000 0100 000
10	0000 1011	26	0000 0011 111
11	0000 1010	27	0000 0011 110
12	0000 1001	28	0000 0011 101
13	0000 1000	29	0000 0011 100
14	0000 0111	30	0000 0011 011
15	0000 0110	31	0000 0011 010
16	0000 0101 11	32	0000 0011 001
		33	0000 0011 000
		MBA Stuffing	0000 0001 111
		Start code	0000 0000 0000 0001

Type Information (MTYPE):
Variable Length

Variable length codewords give information about the macroblock and which data elements are present. Macroblock types, included elements and VLC words are listed in Table 2.

MTYPE is always included in transmitted macroblocks.

Table 2
VLC table for MTYPE

Prediction	MQUANT	MVD	CBP	TCOEFF	VLC
Intra				x	0001
Intra	x			x	0000 001
Inter			x	x	1
Inter	x		x	x	0000 1
Inter + MC		x			0000 0000 1
Inter + MC		x	x	x	0000 0001
Inter + MC	x	x	x	x	0000 0000 01
Inter + MC + FIL		x			001
Inter + MC + FIL		x	x	x	01
Inter + MC + FIL	x	x	x	x	0000 01

Note 1: 'x' means that the item is present in the macroblock
Note 2: It is possible to apply the filter in a non-motion compensated macroblock by declaring it as MC + FIL but with a zero vector.

Quantizer (MQUANT): 5 bits

MQUANT is present only if so indicated by MTYPE.

A codeword of 5 bits signifies the quantizer to be used for this and any following blocks in the group of blocks until overridden by any subsequent MQUANT.

Codewords for MQUANT are the same as for GQUANT.

Motion Vector Data (MVD): Variable length

Motion Vector Data is included for all MC macroblocks. MVD is obtained from the macroblock vector by subtracting the vector of the preceding macroblock. For this calculation the vector of the preceding macroblock is regarded as zero in the following three situations:

1) Evaluating MVD for macroblocks 1, 12 and 23.
2) Evaluating MVD for macroblocks in which MBA does not represent a difference of 1.
3) MTYPE of the previous macroblock was not MC.

MVD consists of a variable length codeword for the horizontal component followed by a variable length codeword for the vertical component. Variable length codes are given in Table 3.

Advantage is taken of the fact that the range of motion vector values is constrained. Each VLC word represents a pair of difference values. Only one of the pair will yield a macroblock vector falling within the permitted range.

Coded Block Pattern (CBP): Variable length

CBP is present if indicated by MTYPE. The codeword gives a pattern number signifying those blocks in the macroblock for which at least one transform coefficient is transmitted. The pattern number is given by

$$32*P_1 + 16*P_2 + 8*P_3 + 4*P_4 + 2*P_5 + P_6$$

where Pn is 1 if any coefficient is present for block n, else 0. Block numbering is given in Figure 10.

The codewords for CBP are given in Table 4.

Table 3
VLC table for MVD

MVD	CODE
-16 & 16	0000 0011 001
-15 & 17	0000 0011 011
-14 & 18	0000 0011 101
-13 & 19	0000 0011 111
-12 & 20	0000 0100 001
-11 & 21	0000 0100 011
-10 & 22	0000 0100 11
-9 & 23	0000 0101 01
-8 & 24	0000 0101 11
-7 & 25	0000 0111
-6 & 26	0000 1001
-5 & 27	0000 1011
-4 & 28	0000 111
-3 & 29	0001 1
-2 & 30	0011
-1	011
0	1
1	010
2 & -30	0010
3 & -29	0001 0
4 & -28	0000 110
5 & -27	0000 1010
6 & -26	0000 1000
7 & -25	0000 0110
8 & -24	0000 0101 10
9 & -23	0000 0101 00
10 & -22	0000 0100 10
11 & -21	0000 0100 010
12 & -20	0000 0100 000
13 & -19	0000 0011 110
14 & -18	0000 0011 100
15 & -17	0000 0011 010

Table 4
VLC table for CBP

CBP	CODE	CBP	CODE
60	111	35	0001 1100
4	1101	13	0001 1011
8	1100	49	0001 1010
16	1011	21	0001 1001
32	1010	41	0001 1000
12	1001 1	14	0001 0111
48	1001 0	50	0001 0110
20	1000 1	22	0001 0101
40	1000 0	42	0001 0100
28	0111 1	15	0001 0011
44	0111 0	51	0001 0010
52	0110 1	23	0001 0001
56	0110 0	43	0001 0000
1	0101 1	25	0000 1111
61	0101 0	37	0000 1110
2	0100 1	26	0000 1101
62	0100 0	38	0000 1100
24	0011 11	29	0000 1011
36	0011 10	45	0000 1010
3	0011 01	53	0000 1001
63	0011 00	57	0000 1000
5	0010 111	30	0000 0111
9	0010 110	46	0000 0110
17	0010 101	54	0000 0101
33	0010 100	58	0000 0100
6	0010 011	31	0000 0011 1
10	0010 010	47	0000 0011 0
18	0010 001	55	0000 0010 1
34	0010 000	59	0000 0010 0
7	0001 1111	27	0000 0001 1
11	0001 1110	39	0000 0001 0
19	0001 1101		

8.2.4 Block Layer. A macroblock comprises four luminance blocks and one of each of the two color difference blocks (Figure 10).

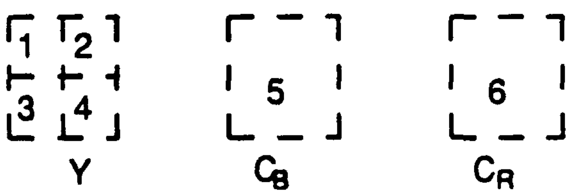


Figure 10
Arrangement of Blocks in a Macroblock

Data for a block consists of codewords for transform coefficients followed by an end of block marker (Figure 11). The order of block transmission is as in Figure 10.

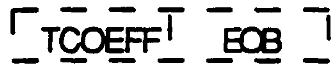


Figure 11
Structure of Block layer

Transform Coefficients (TCOEFF)

Transform coefficients data is always present for all 6 blocks in a macroblock when MTYPE indicates INTRA. In other cases MTYPE and CBP signal which blocks have coefficient data transmitted for them. The quantized transform coefficients are sequentially transmitted according to the sequence given in Figure 12.

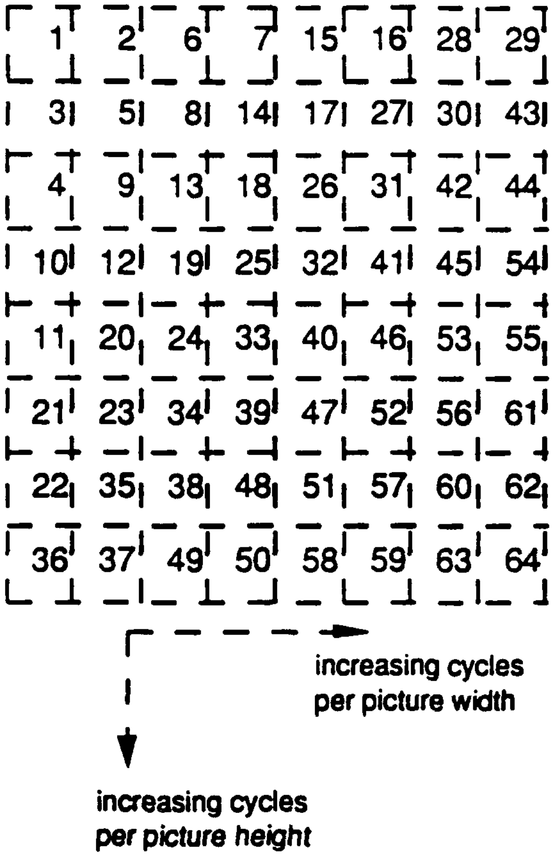


Figure 12
Transmission order for transform coefficients

The most commonly occurring combinations of successive zeros (RUN) and the following value (LEVEL) are encoded with variable length codes. Other combinations of (RUN,LEVEL) are encoded with a 20 bit word consisting of 6 bits ESCAPE, 6 bits RUN and 8 bits LEVEL. For the variable length encoding there are two code tables, one being used for the first transmitted LEVEL in INTER, INTER+MC and INTER+MC+FL blocks, the second for all other LEVELS except the first one in INTRA blocks which is fixed length coded with 8 bits.

Codes are given in Table 5.

Table 5
VLC table For TCOEFF

The most commonly occurring combinations of zero-run and the following value are encoded with variable length codes as listed in the table below. End of Block (EOB) is in this set. Because CBP indicates those blocks with no coefficient data, EOB cannot occur as the first coefficient. Hence EOB can be removed from the VLC table for the first coefficient.

The last bit 's' denotes the sign of the level, '0' for positive
'1' for negative.

RUN	|LEVEL|	CODE
EOB		10
0	1	1s IF FIRST COEFFICIENT IN BLOCK (Note - Never used in INTRA macroblocks)
0	1	11s NOT FIRST COEFFICIENT IN BLOCK
0	2	0100 s
0	3	0010 1s
0	4	0000 110s

0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1010 1s
2	1	0101 s
2	2	0000 100s
2	3	0000 0010 11s
2	4	0000 0001 0100 s
2	5	0000 0000 1010 0s
3	1	0011 1s
3	2	0010 0100 s
3	3	0000 0001 1100 s
3	4	0000 0000 1001 1s
4	1	0011 0s
4	2	0000 0011 11s
4	3	0000 0001 0010 s
5	1	0001 11s
5	2	0000 0010 01s
5	3	0000 0000 1001 0s
6	1	0001 01s
6	2	0000 0001 1110 s
7	1	0001 00s
7	2	0000 0001 0101 s
8	1	0000 111s
8	2	0000 0001 0001 s
9	1	0000 101s
9	2	0000 0000 1000 1s
10	1	0010 0111 s
10	2	0000 0000 1000 0s
11	1	0010 0011 s
12	1	0010 0010 s
13	1	0010 0000 s

14	1	0000 0011 10s
15	1	0000 0011 01s
16	1	0000 0010 00s
17	1	0000 0001 1111 s
18	1	0000 0001 1010 s
19	1	0000 0001 1001 s
20	1	0000 0001 0111 s
21	1	0000 0001 0110 s
22	1	0000 0000 1111 1s
23	1	0000 0000 1111 0s
24	1	0000 0000 1110 1s
25	1	0000 0000 1110 0s
26	1	0000 0000 1101 1s
ESCAPE		0000 01

The remaining combinations of (RUN, LEVEL) are encoded with a 20 bit word* consisting of 6 bits ESCAPE, 6 bits RUN and 8 bits LEVEL.

RUN is a 6 bit fixed length code. LEVEL is a 8 bit fixed length code.

RUN	CODE
0	0000 00
1	0000 01
2	0000 10
.	.
.	.
63	1111 11

LEVEL	CODE
-128	FORBIDDEN
-127	1000 0001
.	.
.	.
.	.
-2	1111 1110
-1	1111 1111
0	FORBIDDEN
1	0000 0001
2	0000 0010
.	.
127	0111 1111

*<u>Note</u> - Use of this 20 bit word form for encoding the combinations listed in the VLC table is not prohibited.

For all coefficients other than the INTRA DC one the reconstruction levels (REC) are in the range -2048 to 2047 and are given by clipping the results of the following formulae;

$$\left.\begin{array}{ll} REC = QUANT*(2*LEVEL+1) & ;LEVEL>0 \\ REC = QUANT*(2*LEVEL-1) & ;LEVEL<0 \end{array}\right\} QUANT = \text{"odd"}$$

$$\left.\begin{array}{ll} REC = QUANT*(2*LEVEL+1)-1 & ;LEVEL>0 \\ REC = QUANT*(2*LEVEL-1)+1 & ;LEVEL<0 \end{array}\right\} QUANT = \text{"even"}$$

$$REC = 0 \;;LEVEL = 0$$

Note: QUANT ranges from 1 to 31 and is transmitted by either GQUANT or MQUANT.

Reconstruction levels (REC)

	QUANT												
LEVEL	1	2	3	4	.	8	9	.	17	18	.	30	31
-127	-255	-509	-765	-1019	.	-2039	-2048	.	-2048	-2048	.	-2048	-2048
-126	-253	-505	-759	-1011	.	-2023	-2048	.	-2048	-2048	.	-2048	-2048
.	.	.	.	.	.	.	.	.	.	.	.	.	.
-2	-5	-9	-15	-19	.	-39	-45	.	-85	-89	.	-149	-155
-1	-3	-5	-9	-11	.	-23	-27	.	-51	-53	.	-89	-93
0	0	0	0	0	.	0	0	.	0	0	.	0	0
1	3	5	9	11	.	23	27	.	51	53	.	89	93
2	5	9	15	19	.	39	45	.	85	89	.	149	155
3	7	13	21	27	.	55	63	.	119	125	.	209	217
4	9	17	27	35	.	71	81	.	153	161	.	269	279
5	11	21	33	43	.	87	99	.	187	197	.	329	341
.	.	.	.	.	.	.	.	.	.	.	.	.	.
56	113	225	339	451	.	903	1017	.	1921	2033	.	2047	2047
57	115	229	345	459	.	919	1035	.	1955	2047	.	2047	2047
58	117	233	351	467	.	935	1053	.	1989	2047	.	2047	2047
59	119	237	357	475	.	951	1071	.	2023	2047	.	2047	2047
60	121	241	363	483	.	967	1089	.	2047	2047	.	2047	2047
.	.	.	.	.	.	.	.	.	.	.	.	.	.
125	251	501	753	1003	.	2007	2047	.	2047	2047	.	2047	2047
126	253	505	759	1011	.	2023	2047	.	2047	2047	.	2047	2047
127	255	509	765	1019	.	2039	2047	.	2047	2047	.	2047	2047

Note: Reconstruction levels are symmetrical with respect to the sign of LEVEL except for 2047/-2048.

For INTRA blocks the first coefficient is nominally the transform DC value linearly quantized with a stepsize of 8 and no dead-zone. The resulting values are represented with 8 bits. A nominally black block will give 0001 0000 and a nominally white one 1110 1011. The code 0000 0000 is not used. The code 1000 0000 is not used, the reconstruction level of 1024 being coded as 1111 1111 (See Table 6).

Coefficients after the last non-zero one are not transmitted. EOB (End of Block code) is always the last item in blocks for which coefficients are transmitted.

Table 6
Reconstruction levels for INTRA-mode DC coefficient

FLC		Reconstruction level into inverse transform
0000 0001	(1)	8
0000 0010	(2)	16
0000 0011	(3)	24
.		.
.		.
0111 1111	(127)	1016
1111 1111	(255)	1024
1000 0001	(129)	1032
.		.
.		.
1111 1101	(253)	2024
1111 1110	(254)	2032

Note: The decoded value corresponding to FLC 'n' is 8n except FLC 255 gives 1024.

8.3 Multipoint Considerations. The following facilities are provided to support switched multipoint operation.

8.3.1 Freeze Picture Request. Causes the decoder to freeze its displayed picture until a freeze picture release signal is received or a timeout period of at least 6 seconds has expired. The transmission of this signal is via external means (for example by H.221).

8.3.2 Fast Update Request. Causes the encoder to encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this signal is via external means (for example by H.221).

8.3.3 Freeze Picture Release. A signal from an encoder which has responded to a Fast Update Request and allows a decoder to exit from its freeze picture mode and display decoded pictures in the normal manner. This signal is transmitted by Bit 3 of PTYPE (see 8.2.1) in the Picture Header of the first picture coded in response to the Fast Update Request.

9. Transmission Coder

9.1 Bit Rate. The transmission clock is provided externally (for example from an I.420 interface).

9.2 Video Data Buffering. The encoder must control its output bitstream to comply with the requirements of the Hypothetical Reference Decoder defined in section 11.

When operating with CIF the number of bits created by coding any single picture must not exceed 256 Kbits. K=1024.

When operating with QCIF the number of bits created by coding any single picture must not exceed 64 Kbits.

In both the above cases the bit count includes the Picture Start Code and all other data related to that picture including PSPARE, GSPARE and MBA Stuffing. The bit count does not include error correction framing bits, fill indicator (Fi), fill bits or error correction parity information described in 9.4 below.

Video data must be provided on every valid clock cycle. This can be ensured by the use of either the fill bit indicator (Fi) and subsequent fill all 1's bits in the error corrector block framing (See Figure 13) or MBA Stuffing (see 8.2.3).

9.3 Video Coding Delay. This item is included in this recommendation because the video encoder and video decoder delays need to be known to allow audio compensation delays to be fixed when video is used to form part of a conversational service.

This will allow lip synchronization to be maintained. A method by which the delay figures are established is recommended in appendix 1. Other delay measurement methods may be used but they must be designed in a way to produce similar results to the method given in appendix 1.

9.4 Forward Error Correction for Coded Video Signal

9.4.1 Error correcting code. The transmitted bit-stream shall contain a BCH(511,493) Forward Error Correction Code. Use of this by the decoder is optional, but will be considered in evaluating the price-performance characteristics of competing vendor proposals.

9.4.2 Generator Polynomial

$$g(x) = (x^9+x^4+1)(x^9+x^6+x^4+x^3+1)$$

Example: For the input data of '01111...11' (493 bits) the resulting correction parity bits are '011011010100011011' (18 bits).

9.4.3 Error Correction Framing. To allow the video data and error correction parity information to be identified by a decoder an error correction framing pattern is included. This consists of a multiframe of 8 frames, each frame comprising 1 bit framing, 1 bit fill indicator (Fi), 492 bits of coded data (or fill all 1's) and 18 bits parity. The frame alignment pattern is

$$(S_1S_2S_3S_4S_5S_6S_7S_8) = (00011011)$$

See Figure 13 for the frame arrangement. The parity is calculated against the 493-bits including Fill Indicator (Fi).

The fill indicator (Fi) can be set to zero by an encoder. In this case only 492 consecutive fill bits (fill all 1's) plus parity are sent and no coded data is transmitted. This may be used to meet the requirement in 9.2 to provide video data on every valid clock cycle.

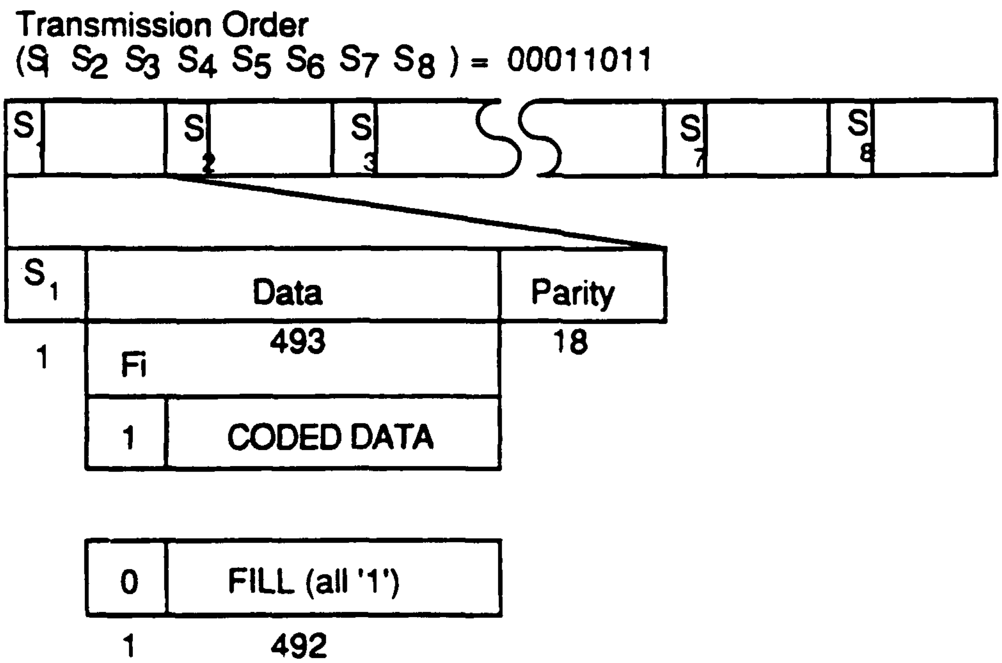


Figure 13
Error correcting frame

9.4.4 Re-Lock Time for Error Corrector Framing. Three consecutive error correction framing sequences (24 bits) should be received before frame lock is deemed to have been achieved. The decoder should be designed such that frame lock will be reestablished within 34000 bits after an error corrector framing phase change.

Note: This assumes that the video data does not contain 3 correctly phased emulations of the error correction framing sequence during the re-locking period.

10. Inverse Transform Accuracy Specification

(1) Generate random integer pel data values in the range -L to +H according to the random number generator given below ('C' version). Arrange into 8 by 8 blocks. Data sets of 10,000 blocks should each be generated for (L=256, H=255), (L=H=5) and (L=H=300).

(2) For each 8 by 8 block, perform a separable, orthonormal, matrix multiply, Forward Discrete Cosine Transform according to the following transfer function using at least 64-bit floating point accuracy.

$$f(u,v) = 1/4\, C(u)\, C(v) \sum_{x=0}^{7} \sum_{y=0}^{7} F(x,y) \cos[Pi(2x+1)u/16] \cos[Pi(2y+1)v/16]$$

with u, v, x, y = 0, 1, 2, ..., 7

where:
x,y = spacial coordinates in the pel domain
u,v = coordinates in the transform domain

C(u) = 1/SQRT(2) for u = 0, otherwise 1.
C(v) = 1/SQRT(2) for v = 0, otherwise 1.

(3) For each block, round the 64 resulting transformed coefficients to the nearest integer values. Then clip them to the range -2048 to +2047. This is the 12-bit input data to the inverse transform.

(4) For each 8 by 8 block of 12-bit data produced by step 3, perform a separable, orthonormal, matrix multiply, Inverse Discrete Transform (IDCT) using at least 64-bit floating point accuracy. Round the resulting pels to the nearest integer and clip to the range -256 to +255. These blocks of 8 by 8 pels are the "reference" IDCT output data.

(5) For each 8 by 8 block produced by step 3, apply the IDCT under test and clip the output to the range -256 to +255. These blocks of 8 by 8 pels are the "test" IDCT output data.

(6) For each of the 64 IDCT output pels, and for each of the 10,000 block data sets generated above, measure the peak, mean and mean square error between the "reference" and the "test" data.

(7) For any pel, the peak error should not exceed 1 in magnitude. For any pel, the mean square error should not exceed 0.06. Overall, the mean square error should exceed 0.02. For any pel, the mean mean error should not exceed 0.015 in

magnitude. Overall, the mean error should not exceed 0.0015 in magnitude.

(8) All zeros in must produce all zeros out.

(9) Rerun the measurements using exactly the same data values of step 1, but change the sign on each pel.

'C' Program for random number generation

```
/* L and H must be long, that is 32 bits */
long rand(L,H)
long      L,H
{
    static long randx = 1;        /* long is 32 bits */
    static double z = (double)0x7fffffff;

    long i,j;
    double x;                     /* double is 64 bits */

    randx = (randx * 1103515425) + 12345;
    i = randx & 0x7ffffffe;       /* keep 30 bits */
    x = ( (double)i ) / z;        /* range 0 to 0.99999... */
    x *= (L+H+1);                 /* range 0 to < L+H+1 */
    j = x;                        /* truncate to integer */
    return( j - L);               /* range -L to H */
    }
```

11. Hypothetical Reference Decoder

The Hypothetical Reference Decoder (HRD) is defined as follows:

(1) The HRD and the encoder have the same clock frequency as well as the same CIF rate, and are operated synchronously.

(2) The HRD receiving buffer size is (B + 256 Kbits). The value of B is defined as follows:

$B = 4R_{max}/29.97$

where R_{max} is the maximum video bit rate to be used in the connection.

(3) The HRD buffer is initially empty.

(4) The HRD buffer is examined at CIF intervals (±33ms). If at least one complete coded picture is in the buffer then all the data for the earliest picture is instantaneously removed (e.g. at t_{n+1} in Figure 14 below). Immediately after removing the above data the buffer occupancy must be less than B. This is a requirement on the coder output bitstream including coded picture data and MBA stuffing but not error correction framing bits, fill indicator (Fi), fill bits or error correction parity information described in 9.4.

To meet this requirement the number of bits for the (N+1)th coded picture d_{N+1} must satisfy:

$$d_{N+1} \geq b_N + \int_{t_N}^{t_{N+1}} R(t)dt - B$$

where b_N is buffer occupancy just after the time t_N, t_N is the time the Nth coded picture is removed from the HRD buffer, R(t) is the video bit rate at the time t.

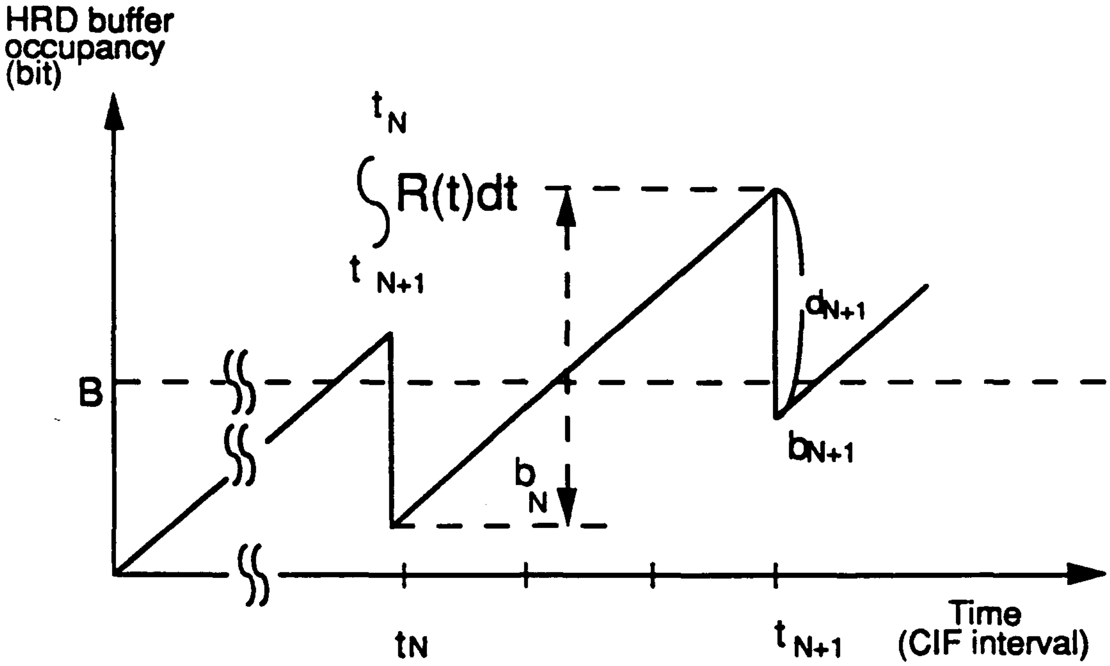


Figure 14
HRD buffer occupancy

Note: Time (t_{N+1} - t_N) is an integer number of CIF picture periods (1/29.97, 2/29.97, 3/29.97, ...).

Appendix 1 Codec Delay Measurement Method.

The video encoder and video decoder delays will vary depending on implementation. The delay will also depend on the picture format (QCIF, CIF) and data rate in use. This section specifies the method by which the delay figures are established for a particular design. To allow correct audio delay compensation the overall video delay needs to be established from a user perception point of view under typical viewing conditions.

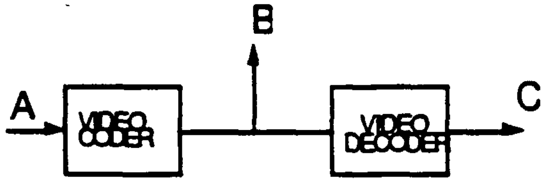


Figure 15
Measuring points

Point A is the video input to the video coder. Point B is the channel output from the video terminal (i.e. including any FEC, channel framing etc.) Point C is the video output from the decoder.

A video sequence lasting more than 100 seconds is connected to the video coder input (point A) in Figure 15 above. The video sequence should have the following characteristics.

(1) It should contain a typical moving scene consistent with the intended purpose of the video codec.

(2) It should produce a minimum coded picture rate of 9.5 Hz at the bit rate in use.

(3) It should contain a visible identification mark at intervals through out the length of the sequence. The visible identification should change every 97 video input frames and be located within the picture area represented by the first GOB in the picture. For example the first block in the picture could change from black to white at intervals of 97 video frame periods. The identification mark should be chosen so that it can be detected at point B and does not significantly contribute to the overall coding performance.

The codec and video sequence should be arranged so that the bit stream contains less than 10% stuffing (MBA stuffing + error correction fill bits).

The encoder delay is obtained by measuring the time from when the visible identification changes at point A to the time that the change is detected at point B. Similarly, the decoder delay is obtained by taking measurements at points B and C.

Several measurements should be made during the sequence length and the average period obtained. Several tests should be made to ensure that a consistent average figure can be obtained for both encoder and decoder delay times.

Average results should be obtained for each combination of picture format and bit rate within the capability of the particular codec design.

Note: Due to pre and post temporal processing it may be necessary to take a mid-level for establishing the transition of the identification mark at points B and C.

APPENDIX C

DRAFT CCITT RECOMMENDATIONS FOR MULTIPOINT AUDIOVISUAL SERVICES

T1Y1.1/91-004

International Telegraph and Telephone
Consultative Committee

COM XV-WD4/1
Issue: 12 October 1990

Period 1989-92

Original: English

Question 4/XV

STUDY GROUP XV - CONTRIBUTION WD4/1

Title: **MULTIPOINT AUDIOVISUAL SERVICES**

Source: Special Rapporteur for Question 4/XV

This contribution concerns four draft Recommendations:

AV.231: Multipoint Control Unit for Audiovisual Services

AV.243: System for establishing communication between three or more audiovisual terminals using digital channels up to 2Mbit/s

AV.441: Call-control procedures for real-time audiovisual conference calls - procedures not requiring special network capabilities

AV.442: Call-control procedures for real-time audiovisual conference calls - procedures requiring special network capabilities (responsibility for this lies with SG XI)

In the case of the last two items, some requirements are proposed as a starting point for the work.

DRAFT RECOMMENDATION AV.231: MULTIPOINT CONTROL UNIT FOR AUDIOVISUAL SYSTEMS

CONTENTS

1. SCOPE

2. FUNCTIONAL REPRESENTATION

A multipoint call may be represented as in Figure 1, wherein are shown a number of terminals T, not necessarily identical, linked individually into a network by symmetrical bidirectional digital connections, not necessarily all of the same capacity. There is no particular limit set by the system to the number N of terminals connected in the call, though in practice, depending on implementation, the difficulties and cost will rise as N increases while performance tends to fall.

In the representation of Figure 1, the network need only be described by the signal flows at its ports, and their interdependencies. The hardware realisation need not be of concern: there may be a single MCU at one location; alternatively the functions may be distributed to two or more locations, but in practical terms we then refer to a series of single MCUs linked together in a chain. In this Recommendation, the text applies in general to both single-location and distributed MCUs, and the linking of MCUs is only treated specifically where there is a particular need to do so.

The MCU is represented in more detail in Figure 2.

Each port of the MCU has a Network Interface Unit, with associated call control; on the MCU side of the Network Interface Unit, the signal flows are contained in one or more bidirectional channels of equal capacity, according to the transfer rates listed in Rec H.221, Annex 2. The incoming flow is passed to the Demultiplexer, which extracts the several types of information (audio, video, data, and control) and passes them to their respective processors. The processors are controlled in such a way that an appropriate output from each is made available for transmission to every terminal; the latter are brought together in the Multiplexer to be combined into the outgoing channels.

The call-control units and processor are outside the scope of this Recommendation (see Rec AV.~~400~~ 440); the other units are described in the following sections.

3. FUNCTIONAL UNITS

3.1 Network Interface Unit

3.2 Demultiplexer

The signal entering the demultiplexer is that transmitted by a terminal fully conforming to Rec H.221, so the operation is analogous to that of the receiving side of a terminal, namely:

- recovery of frame and multiframe alignment
- buffering, synchronisation and ordering of multiple channels if relevant
- extraction of BAS codes and forwarding some of them to the control processor
- extraction of encryption vectors and decryption if relevant
- extraction of audio and forwarding to the audio processor
- extraction of video and forwarding to the video processor
- extraction of data and forwarding to the data processor

Correct timing relationships must be maintained between mode-control BAS and the related audio, video, and data.

3.3 Audio Processor Unit (APU)

The audio processor prepares N audio outputs r_j from N audio inputs s_i, by switching, mixing, or a combination of these. Mixing requires the addition of linear signals S_i obtained by decoding s_i to linear (PCM or analogue), and the recoding of the responses R_j to appropriate transmission formats r_j.

An audio-mixing MCU generally creates the responses

$$R_j = \sum_{i=0}^{N} \mu_i S_i \qquad \text{where } \mu_i = \delta_i \text{ for } i=j, \quad \mu_i = 1 \text{ otherwise}$$

The value of δ_i must be very small, as it represents an echo to the originating terminal T_i: the relative magnitude of this echo must be less than -45dB(?) if there is low bit rate moving video in the call, and less than -35dB(?) otherwise.

Under certain circumstances, two terminals T_x and T_y may be removed from the mixing function and interconnected separately:

$$R_x = S_y, \; R_y = S_x, \; R_j = \sum \mu_i S_i \quad \text{where x and y are excluded from the summation}$$

It may be appropriate to limit the summation in other ways. For example, if N is large, the background noises from all terminals may sum to an annoying level: the control, monitoring the magnitudes of S_i, may be set to include in the summation only a limited number of signals whose present or recent past values exceed(ed) a certain threshold. Alternatively, a person may control the summation directly (see section 3.5.2) so that only certain terminals may be heard.

If in either of the above cases the number is limited to one, the MCU becomes audio-switching instead of audio mixing. The manually controlled audio-switched approach is useful where encrypted audio cannot be decrypted at the MCU. The audio unit may also contain a voice synthesiser or recorded message store, able to be connected into the mixing unit or separately to any terminal.

If there is a "mixing" delay in the VPU (see below) greater than that in the APU by more than 30ms, a compensating delay must be added in at the appropriate APU outputs.

3.4 Video Processor Unit (VPU)

The video processor can operate in ways entirely analogous to those described above for the audio processor. "Mixing" takes the form of spatial multiplexing into a split-screen format: to do this the incoming video signals may need a certain amount of preprocessing.

Since the video mixing function is highly complex, the alternative of video switching may be preferred. As for audio switching, the choice of video may be automatic, such that the present speaker (largest value of s_i) receives the picture of the previous speaker, while all other terminals receive the picture of the present speaker; a time delay is incorporated into the switching (value 2s) to avoid excessively frequent image changes, caused by spurious sounds such as coughing, knocking a microphone, and so on.

Again, the video switching may be controlled directly by a person making his own decisions as to which picture is most appropriate.

If there is a codec delay in the APU greater than that in the VPU by more than 60ms, a compensating delay must be added in at the VPU outputs.

3.5 Data Processor (DPU)

Two cases are distinguished, according to whether or not the MCU has equipment for handling the multilayer protocol (MLP) defined in Rec AV.270.

3.5.1 Basic MCU, Without MLP Capability

In this case, only one* data input can be accepted at any one time, any data subsequently arriving at another input being ignored. The data is broadcast to all other outputs, or at least to those outputs determined by the control processor. (*Two, if the connections are at 2B or greater rates)

3.5.2 Enhanced MCU, Having MLP Capability

In this case the data processor can perform both the operation described in Section 3.5.1, and the MLP process described here. The former is applied to data transmitted with MLP "off", and to OLSD (see Rec H.242) where MLP is "on".

The data contained within the MLP is passed to the MLP Processor, which performs one or more of the following functions (see Recs AV.270 ff):

- processing of telematic information

- processing of conference control signals (request/grant floor, chairman token control, audio/video switching).

3.6 Control Processor Unit (CPU)

The control processor is responsible for determining the correct routeing, mixing/switching, format and timing of the audio, video, data and control signals passed to each multiplexer for outward transmission.

3.6.1 Incoming BAS Commands

The CPU ensures that the correct audio decoding algorithm is used on each input to the audio mixer; that data is sent to the broadcast unit or MLP Processor as appropriate.

3.6.2 Outgoing BAS Commands

The CPU ensures that the correct audio encoding algorithm is used on each output from the audio mixer, and that the desired switching or summation has been performed in each case; that the desired switching (or mixing of video signals) has been made to each output of the VPU. It transmits VCF (see Rec H.230) to all relevant terminals at a set time before switching the video sent to them, and VCU to a terminal whose video is about to be sent to another terminal; the procedure for this is set out in section 4.

The CPU switches mode on outgoing streams to accommodate the appropriate combination of audio, video, and data, according to the declared capabilities of the connected terminals (see Rec AV.243).

3.6.3 Incoming BAS Capabilities

The capability codes from all N terminals are stored; whenever a new set is sent by a terminal, it replaces completely the previous set (exception: encryption capability).

3.6.4 Outgoing BAS Capabilities

The values to be sent at each of the N ports are determined according to Rec. AV.243.

3.7 Multiplexer

The multiplexer sets up a frame structure on the outgoing channel(s) according to Rec H.221, and loads into this the BAS values from the CPU and the outputs of the APU, VPU and DPU.

4. VIDEO SWITCHING

When it is decided within the CPU that terminal A, currently receiving the video signal from terminal B, should instead be sent that from terminal C, the following procedure is used (codes VCF, VCU are specified in Rec H.230).

(a) The MCU transmits VCF to terminal A, and immediately afterwards switches video such that the picture from C is transmitted towards A.

(b) Terminal A receives VCF, and freezes its currently displayed picture; it ignores subsequent decoded video information, but continues to track the error-correction framing, and to monitor Picture Headers for the Picture Freeze Release command.

(c) When incoming video to A changes from B-picture to C-picture, error-correction frame alignment is lost, and will take a time T to recover, dependent on the video bit rate and other factors.

(d) After a time greater than T, the MCU transmits VCU to terminal C.

(e) On receipt of VCU, terminal C sends its next video frame in "fast-update" mode (Rec H.261, Section 4.3.2), together with the Picture Freeze Release command.

(f) On receipt of the Picture Freeze Release command, terminal A reverts to displaying the incoming decoded picture.

Note: users at other terminals which have been receiving picture C continuously during the above procedure will nevertheless be aware of the switching action because of the use of the fast-update mode: this is the transmission of a single new picture over a period inversely proportional to video bit rate - at 320 kbit/s this period is likely to be about 0.5 seconds.

5. SUMMARY OF MCU CAPABILITIES

The MCU capabilities must be such as to handle the signals of the terminals with which it is to be used, as listed and defined in Rec H.242.

Note: this section is concerned with the internal capabilities of the MCU; the BAS-capabilities declared at any time on a particular port of an MCU are a combinatorial function of the terminals connected - see Rec AV.243.

(a) Audio: an audio-mixing MCU must possess decoding and encoding capabilities from the set A-0, A-1, A-2, A-3. An audio-switching MCU does not decode any audio signals; internally generated messages may be transmitted as PCM.

(b) Video: an MCU may or may not be able to handle video. If it does so by switching, the different video capabilities defined in Rec H.242 are of no concern; however a video-mixing MCU would have to take them into account.

(c) Transfer Rate: the same values as defined in Rec H.242.

(d) Restricted-Network Capability: for further study.

(e) Data (except MLP): since the only function of the MCU is broadcasting of data streams, it may be presumed that, possessing this capability at all, it is available at rates up to the highest transfer rate.

(f) MLP: the MCU requires a considerable body of software to handle MLP; however, while there might well be speed limitations arising from computing power, it would be unwise to impose a physical limit lower than the highest transfer rate.

(g) Encryption: for further study.

(h) VLE Capability: the MCU may or may not be able to handle variable length extension BAS codes.

Examples

(i) A basic ISDN MCU might well possess the following capabilities:

/A-2 + A-3, switched video, T-1B + T-2B, Data (broadcast)/

(ii) An audiographic MCU might be:

/A-2, T-1B, Data, MLP, VLE/

(iii) A videoconference MCU might be:

/A-2, switched video, T-2B + T-HO, Data/

NDK/1973

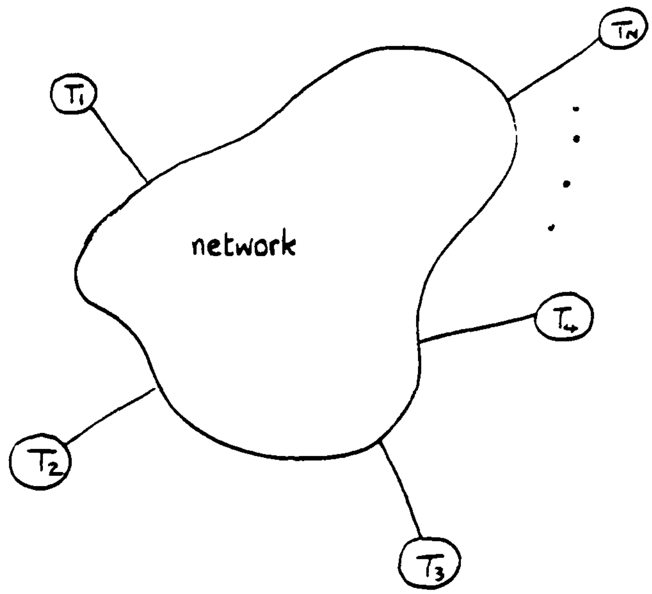


FIGURE 1.

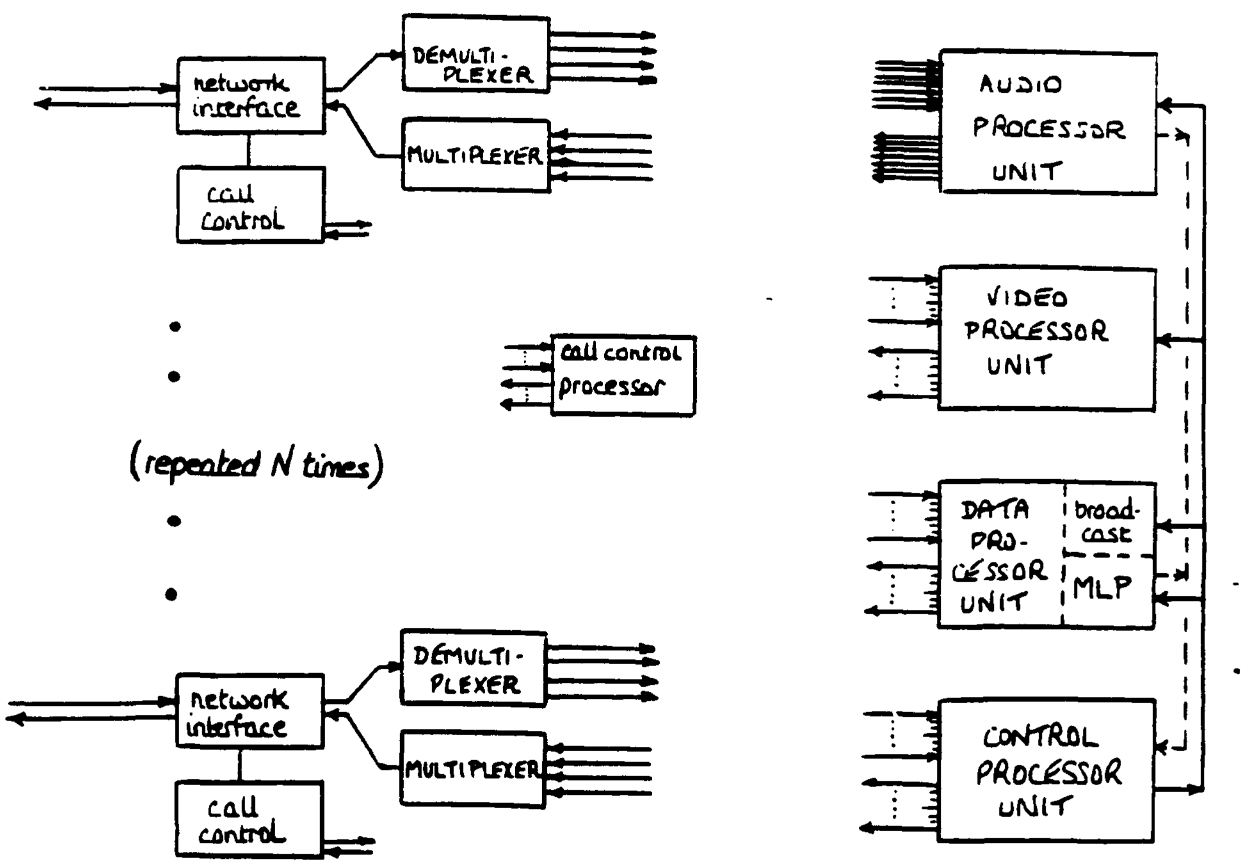


FIGURE 2. Schematic of Multipoint Control Unit

DRAFT RECOMMENDATION AV.243

SYSTEM FOR ESTABLISHING COMMUNICATION BETWEEN THREE OR MORE AUDIOVISUAL TERMINALS USING DIGITAL CHANNELS UP TO 2 MBIT/S

CONTENTS

1. INTRODUCTION AND SCOPE

In a point-to-point connection between two terminals, it is worthwhile to get the best quality working mode, according to the respective capabilities of the terminals; if the two terminals are quite different, it is obvious that the working mode is limited by the lower capability terminal.

In a multipoint connection, it must be assumed that different terminals can be connected to the same multipoint control unit (MCU), in which case it may not be appropriate that the working mode be the ~~lowest~~ highest common one. For example, if three videophones and one digital telephone are connected together via an MCU, it would not be very sensible that all terminals work in an audio mode only. The objective of this paper is to describe a MCU rule to avoid such a situation.

Recommendation AV.242 provides for communication between two audiovisual terminals connected point-to-point, using the frame structure defined in Rec H.221 and the control and indication symbols defined in AV.230.

Three or more audiovisual terminals may be put into communication to form a conference call, by means of one or more multipoint control units (MCU). The means by which digital channels are established between terminals and MCUs, and between MCUs, is outside the scope of this recommendation, although reference is made to relevant circumstances in section 3. This recommendation concerns only the flows of signals along the fixed digital paths, which may be at 64 kbit/s or multiples thereof up to 2048 kbit/s. The flow consists of a multiplex of audio, video, telematic, user data, and control and indication signals, which must be handled by the MCU in a way which is satisfactory to the users.

The signal multiplex on each path is fully in accordance with H.221: the BAS commands define explicitly how the demultiplexer at the end of each link shall operate. Likewise the basic procedures for initialisation and mode switching are fully in accordance with those defined in AV.242 for point-to-point working. However the composition of the multiplexed signal transmitted by each terminal and by the MCU is determined by terminal procedures and MCU procedures, as follows:

(a) terminal procedures are defined in service-specific system recommendations, such as AV.320 for visual telephony;

(b) MCU procedures are defined in this recommendation, and are not of themselves service-specific;

(c) multi-layer protocol (MLP): by making use of the MLP defined in the AV.270 series, MCU and terminal procedures may be greatly enhanced, offering far more sophisticated specific applications to the user.

2. DEFINITIONS

Multipoint Control Unit (MCU):

Convenor Terminal: a terminal possessing a token conveying a certain measure of authority over the operation of the MCU; the token may be assigned by prearrangement, by an operator, or by protocol during the call.

Convenor Port: that port of the MCU to which the convenor terminal is connected.

Primary and Secondary Ports: while all ports of an MCU may be physically the same, distinctions may be made by the internal software, on the basis of declared terminal capabilities, such that the ports are not all treated on an equal basis. In general, a multipoint call will involve two or more terminals intercommunicating on an equal basis, at their highest common capability; the MCU would designate as "primary" those ports to which these terminals are connected, and for simplicity these terminals can be referred to as "primary terminals" for the purposes of this one call. The selection of an appropriate common level for primary communication is specified in section 6.

One or more additional terminals may take part in the multipoint call, even though they do not have sufficient capability to communicate on an equal basis with primary terminals; these may be designated "secondary terminals", communicating with the others only by such compatible signals as can be made available (eg, speech only), the MCU having designated the appropriate port accordingly. Note that if this provision were not made, then the addition of a telephony terminal to a videophone conference would cause all picture transmission to be discontinued.

The concept of primary/secondary terminals is introduced in this Recommendation for clarity in the description of procedures. A more generalised MCU may be devised in which a rigid distinction between primary and secondary ports is not necessary: this matter is discussed further in Appendix 1.

3. MULTIPOINT CONFIGURATIONS

Star: all terminals connected to a single MCU; all primary terminals are connected at the same effective bit rate, being 64 kbit/s or a multiple up to 2048 kbit/s (rates defined in Annex 2 to Recommendation H.221); secondary terminals may be connected at a lower rate.

Dumb-bell: terminals are connected to one of two MCUs, which are themselves interconnected at the same effective rate as the primary terminals.

MCU Chain: three or more MCUs may be connected in tandem (but not in a ring) with terminals connected to each, the MCUs being interconnected at the same effective rate as primary terminals.

Call Set-up Configurations: arrangements for setting up multipoint call are described in Recommendation AV.440.

4. MULTIPOINT CONTROL UNIT CAPABILITIES AND REQUIREMENTS

4.1 MCU Capabilities

Communication between each terminal and an MCU is on the same basis as between two terminals in a point-to-point call, and governed by procedures similar to those in Rec H.242. For this purpose the MCU transmits a set of capabilities, known as "MCU capabilities", to ensure that each terminal does not transmit signal which cannot be decoded at the MCU or passed on other connected terminals. It is to be noted that the capabilities are not necessarily those of the MCU itself (ie, possibly related to other terminals) and not necessarily the same at each port (ie, the terminals may be treated differently), and furthermore will vary during the call (see Sections 6-9).

Audio Capability

Since audio is usually mixed at the MCU (see Rec AV.231), the capability is set at the common mode to be enforced; secondary terminals may be added in without reaching the common "primary" audio mode.

Transfer-rate Capability

This is set to the highest common value for primary terminals, unless the MCU itself has a lower capability determined by its access ports.

Video Capability

This is set to the highest common value for primary terminals; towards secondary terminals it may be advisable to declare no video-capability, if it is thought desirable to avoid the circumstance in which images from the secondary can be seen by primaries, but the secondary terminal cannot see primary images because its receiving capability is too low.

Restricted Network Capability

This is set if the primary communication is chosen to be in modes of the restricted type (see Rec H.221, Annex 9).

Other Capabilities (see Rec H.242, Capability Table Z)

These are set towards primary terminals if common to primary terminals; towards secondaries they may be set if the configuration of the MCU is such that it can cope (for example, MLP up to 6.4 kbit/s may be possible to all terminals).

4.2 MCU Requirements

The requirements of MCUs are defined in Rec AV.231; a few points are reproduced below.

Basic MCU: such an MCU has no MLP capability, and no special video/data storage/conversion facilities. Secondary terminals are connected by audio only (for further study).

Enhanced MCU: optional enhancements include MLP capability, and/or the means for conversion of signals.

All MCUs must conform to revisions of Recs H.221 and AV.242 and this Recommendation.

Treatment of Signals:

- audio signals are generally mixed in such a way that every terminal is sent a linear addition of the audio signals from all other terminals;

- moving video of participants, automatic switching based on speech power evaluation, present speaker image distributed to all others; previous speaker image sent to present speaker (alternative transmissions under user control are invoked by means of an MLP or by means of the MCU control);

- data, broadcast from one terminal to all others (alternative transmissions under user control are invoked by means of an MLP).

5. TERMINAL REQUIREMENTS AND OPTIONS

All terminals must conform to the provisions of Recs H.221, H.230 and H.242.

Optional Convenor Functions: request/release of convenor token, using MCR, MCT; call control capability (see Rec AV.110).

6. BASIC MCU: INITIALISATION PROCEDURES

Some terminals may be able to recognise whether they are connected to another terminal or to a MCU, but this is not the general situation. It is therefore necessary to adopt a similar procedure for point-to-point and MCU connections.

At the beginning of the call, the terminal sends its capabilities and is waiting to receive frame structure and capabilities, as described in H.242, with transmission in a mode OF only. Therefore, the MCU must send appropriate capabilities, according to the phase of the multiple call set-up: at the completion of call set-up on each port, the MCU begins transmitting in mode OF with capabilities as indicated below.

6.1 First ~~Call~~ Terminal Connected to MCU

The MCU registers the capability of this first terminal, which it designates T1. If T1 declares MLP-cap, and if also the MCU has MLP handling equipment, MLP-cap is transmitted in both directions and the MLP channel is opened by the MCU, permitting greatly enhanced communication possibilities (such as vetting of subsequent terminals by T1 before their addition to the audio mixer, etc).

If T1 does not declare MLP-cap, the MCU transmits only PCM-audio capability in mode OF, with the synthesised message M1 (see Rec AV.440, section N) and the C&I symbol MIZ (see Rec H.230) indicating that no other terminals are yet connected.

Optionally, if there is mutual video capability this could also be invoked (for further study).

6.2 Second ~~Call~~ Terminal Connected to MCU

The MCU registers the capability of the second terminal upon recovering frame alignment, while transmitting those capabilities of T1 that it can itself handle (this may, for example, exclude MLP or video). Terminals T1 and T2 my then switch to other suitable modes, according to their own procedures.

6.3 Third ~~Call~~ Terminal Connected

MCU registers the capability of terminal T3; using criteria given in section 6.4 below, the MCU determines whether T3 is to be treated as primary or secondary; if T3 is to be primary, a further check is made to determine whether T2 should be dropped to secondary status. The MCU then transmits on all its ports capability equivalent to the common capability of the primary terminals. It also transmits to any secondary terminal the C&I symbol MIS, indicating the secondary status accorded; if possible, the terminal should display its secondary status to its user, though this may become apparent from verbal communication also.

6.4 Selection of Primary Capability

The primary capability may be fixed or variable, according to the chosen one of several approaches, as follows:

(a) specified by the service provider or manufacturer, and fixed in the MCU;

(b) set by the convenor terminal using MLP ~~communication~~ tal-kit;

(c) set automatically by the MCU at the capability of the convenor terminal;

(d) set by the MCU, initially at the capability of the convenor terminal but subsequently lowered to match communication as much as possible.

The algorithm for method (d) is set by the manufacturer, taking into account the following principles if relevant.

(i) If audio is mixed (rather than only switched) in the MCU then since the mixing must be linear no restraint is placed on the speech bandwidth or coding algorithm on any path: if two or more terminals are capable of wideband audio (in addition to video if relevant) they may use this independently of whether they are designated primary or secondary. However if the selection of wideband audio between two or more terminals, together with video, would have the result that another terminal (eg, 13 videophone) could not receive video, then it may be advisable to exclude wideband audio from the primary set.

(ii) If the convenor and at least one other terminal have video capability, then all terminals without video should probably be designated secondary..

When referring to a "higher" or "lower" capability with respect to audio with video only, the hierarchy is taken to be that of Table 1/H.320, together with the principle that all CIF modes are higher than QCIF, regardless of associated minimum picture interval capability. It must not, however, be inferred that a terminal capability includes all lower values: for example, CIF/7.5 does not necessarily include QCIF/30, and A-3 does not necessarily include A-2.

(Annex 2 shows the procedure for an ISDN MCU resulting from the above principles

6.5 Fourth Call Connected

The procedure followed is essentially that of section 6.3 above, establishing whether the newcomer is to be accorded primary or secondary status, and checking whether any existing primary terminals should be dropped to secondary.

6.6 Extension to Multiples of 64 kbit/s

If the convenor and the terminals selected as primary, according to the process of section 6.4 above, have the capability to cope with at least m additional channels, then the transmitted MCU transfer-rate capability reflects the (m + 1) rate to relevant terminals, and the additional channels are set up according to the terminal procedures.

The MCU transmits the symbol MCS, to ensure that it retains control of the moment when expansion of audiovisual signals to higher rates can be released. When all the requested additional channels are available, or after expiry of a timer (5m seconds?), MCN is transmitted, enabling primary terminals to make use of the additional capacity.

Note: the procedure may not be necessary for two-B calls; it may be sufficient for the MCU to allow expansion to two-B as soon as the convenor and one other terminal have two B channels connected, temporarily dropping to secondary status any other terminals which have not yet completed the second B-channel connection.

7. BASIC MCU: MODE SWITCHING FOR DATA DISTRIBUTION IN SIMPLE MULTIPOINT CONFERENCES

In a point-to-point connection, a terminal may transmit data to the other end (having received a data capability BAS) simply by mode switching and putting the data into the appropriate position in the multiplex.

In a multipoint call involving video this will not generally be possible, because not only must the outgoing video rate be dropped but also that of video signals emanating from other terminals.

The terminal T_D wishing to send data opens the appropriate data channel using the BAS data command, but puts no data into it. (The terminal is aware of the multipoint configuration, having received MIC from the MCU.)

Detecting the BAS command change, the MCU transmits MCS to all terminals other than the data originator, enforcing symmetrical transmission, thereby leaving the desired capacity available. The MCU sends the new data command to all terminals. When T_D detects incoming data command change, it is free to transmit the data. The MCU invokes the data-broadcast function from the time all its incoming videostreams have been lowered in bit rate. When T_D ends transmission of data, it closes the appropriate data channel using the BAS data command, the MCU then transmits MCN to all other terminals.

8. ENHANCED MCU WITH MLP: INITIALISATION PROCEDURE

The procedure essentially follows that of section 6, the capability BAS being included in the capability exchanges. When the MCU has detected that one or more terminals have MLP capability, the 6.4 kbit/s data channel is opened to all terminals, and by transmitting the symbol MCS it is ensured that the same capacity is also vacated on the incoming direction to the MCU. It is necessary to do this to all terminals, to ensure that those not having MLP capability do not, for example, transmit too great a video data rate. MLP exchanges may now take place between those terminals having this capability and the MCU, the space capacity to other terminals being left empty.

9. BASIC MCU: SUSPENSION PROCEDURE

10. ENCRYPTION

11. PROCEDURES RELATED TO MAINTENANCE

12. SEQUENCING OF BAS CODES

The principles of Rec H.242, section 12, should be followed, with the additions described below.

The MCU transmits the C&I symbol MIC periodically to all terminals, to ensure that they remain aware of particpation in the multipoint call.

13. MCU-MCU INTERCONNECTIONS

For further study. It may be difficult to achieve satisfactory operation by method (d) of Section 6.4, though other methods may still be applicable.

Add: capability exchange during a call

MCU WITHOUT THE NEED FOR PRIMARY/SECONDARY PORTS

Such an MCU will treat each terminal with a degree of independence, by:

- accepting from terminal Ti only such signals as it may usefully pass on to one or more other terminals; this "acceptance" is controlled by the capability declaration by the MCU towards Ti;

- transmitting to Ti only such signals as Ti can digest, according to the capability declaration by Ti towards the MCU.

The difference from the primary/secondary concept lies in the interpretation and implementation of "one or more other terminals", which takes the place of "all primary terminals".

Some of the problems are illustrated here.

(a) In a call involving three 1B videophones and two 2B+ videophones, the former would not receive pictures from the latter, unless the MCU included a video transcoder.

(b) Even when all video rates are the same a similar problem arises from the range of video capabilities.

(c) Further complications arise with the introduction of data paths which can only be opened to some terminals.

PROCEDURE FOR MCU WITH ISDN BASIC ACCESS ONLY

It is assumed that the MCU can handle 1B or 2B connections on all ports, and that the audio mixer is capable of AV.254, G.722, and G.711 on all ports.

The following types of visual telephone are defined in Rec H.320:

Type Xa: with capabilities A-3 and video up to 46.4 kbit/s;

Type Xb1: with capabilities A-0 and A-3 and video up to 64 kbit/s;

Type Xb2: with capabilities A-2 and A-3 and video up to 64 kbit/s;

Type Xc: with capabilities A-3 and A-2 and video up to 110.4 kbit/s.

In fact it is not absolutely clear whether Type Xc includes A-2 capability, but for completeness let us add:

Type Xc*: with capabilities A-3 (and not A-2) and video up to 110.4 kbit/s

We may also consider the cases of two types not recognised in H.320:

Type Xb3: with capabilities A-2 or A-0 (and not A-3) and video up to 64 kbit/s

Type Xb3*: as type Xb3 but with the additional capability to set the video rate at 46.4 kbit/s (Note 1).

1. If there are not two terminals with video capability, go to 9.

2. If there are only two with video capability, and one is Type Xa, the other type Xb3, go to 9.

3. If half or more of the video terminals can transmit video at 46.4 kbit/s (that is, they are of any video Type except Xb3) go to 5.

4. Since half or more of the video terminals are of Type Xb3, go to Outcome III.

5. If there are any Type Xa involved, go to Outcome II.

6. We are now dealing with 2B terminals only, and must determine which video rate is to be used (110.4 kbit/s or 62.4 kbit/s). If there are any terminals (Xb3 or Xb3*, or even Xb1 or 2 if the video decoder speed is limited to 62.4 kbit/s) which cannot do 110.4 kbit/s then the choice must be 62.4 kbit/s (Outcome III).

7. If all can do 110.4 kbit/s but a majority also have wideband audio capability, the MCU could decide that better audio is more important than better video, and so remains with Outcome III.

8. This leaves Outcome I.

9. There can be no video transmission, so we treat all as audio + data terminals. If the MCU and any terminal(s) have MLP capability the appropriate channel(s) should be opened (even only from one terminal to MCU may serve a useful purpose).

10. If two or more terminals have G.722 audio then they use this (Outcome IV).

11. If two or more terminals declare 2B transfer rate capability then the MCU confirms this rate on those connections (Outcome VI); however it would be better for such declarations only to be made after human or MLP negotiation.

12. Data rate capabilities declared by the MCU should be any common declared rates within the bounds of the "Outcomes" table, preferably negotiated in advance by users or MLP.

In the table of outcomes it can be seen that the Type Xb3 has problems: when other terminals are in the majority, it is excluded from the video communication, and when it is in the majority then any 1B visual telephones are excluded from the video communication. The Type Xb3* has no such problem.

It would therefore be well worth providing for the BAS command mode definition to support this type of audiovisual terminal.

Notes

(1) To achieve a video rate of 46.4 kbit/s in a 2B connection it is necessary to vacate bits 1 and 2 of the additional channel, using another BAS command to be defined. This mode would ensure that such a terminal could interwork with a Type Xa terminal via an MCU, or, for that matter, through a two-port device.

(2) The LSD rates listed are those which do not involve switching the audio and video off, and must be common to at least two terminals, not necessarily both primary. For read "less than or equal to".

(3) Video capabilities (QCIF/CIF and minimum picture interval) are declared by the MCU at the highest common capability (HCC) of the terminals included as primary.

TABLE OF OUTCOMES

The table entries show the capabilities to be declard by the MCU in the directions of the various types of terminal, and the resulting video rate in italics.

	PRIMARY						SECONDARY			
OUT COME	Type	T	A	V	LSD	Exclusion	T	A	V	LSD
Outcome I	Xc,c*	2B	A-3	HCC	<46.4k	n/a	1B	A-3,2,0	No	<46.4k
	Xb1,2	2B	A-3	HCC	<46.4k					
				(110.4k)						
Outcome II	Xa,c,c*	1B	A-3	HCC	<6.4k	Xb3	1B	A-3,2,0	No	<6.4k
	Xb1,2	1B	A-3	HCC	<6.4k					
	Xb3*	2B	A-2	HCC	<6.4k					
				(46.4k)						
Outcome III	Xc	2B	A-2	HCC	<14.4k	Xa	1B	A-3,2,0	No	<14.4k
	Xb*(all)*	2B	A-2	HCC	<14.4k					
	Xc*	2B	A-0	HCC	<14.4k					
				(62.4k)						
Outcome IV	A+D	1B	A-2	No	<14.4k		1B	A-3,0	No	<14.4k
Outcome V	A+D	1B	A-3	No	<46.4k		1B	A-0	No	<46.4k
Outcome VI	A+D	2B	A-2	No	<78.4k		1B	A-3,2,0	No	<64k

CONSIDERATIONS FOR RECOMMENDATION AV.440: CALL-CONTROL PROCEDURES FOR REAL-TIME AUDIOVISUAL CONFERENCE CALLS

CONTENTS

1. INTRODUCTION

Conversational services provide for real-time speech communication between users, aided by various other facilities such as video, telematics and data. Communication may be established between three or more audiovisual terminals by means of one or more multipoint control units (MCU) which handle the processing and distribution of the various information signals in an appropriate way. The operation of an MCU is defined in Rec AV.231, and in AV.243 the procedures for in-band system control after call establishment. This Recommendation defines a number of procedures by which connections to an MCU may be established, abstracting the network and in-band signalling requirements to support such procedures.

2. TYPES OF CONFERENCE CALL

The following types of conference call may be distinguished:

(a) "meet-me" type: all participants dial into a pre-arranged MCU;

(b) "dial-out" type: a convenor dials all other participants;

(c) supplementary service, in which a point-to-point call is diverted through an MCU in order to add in new participants;

(d) operator controlled: a non-participant, such as an employee of the service provider, dials the calls to all terminals and transfers them to the MCU;

(e) other types of conference call may also emerge.

The use of conference calls will not become widespread unless they are easy to establish: the procedures must be based on sound human-factor principles. The procedures described in this Recommendation are for guidance only: more important is the standardisation of the signals which facilitate the various procedures. It is the responsibility of manufacturers and service providers to ensure that presentation of the service to the user is optimised.

Attention is drawn to the need to provide conference call services that can be used from all types of terminal, including simple videophones and audio/data terminals that are not specially equipped for conferencing. It is convenient to refer to a basic conference call between such terminals, as distinct from an enhanced conference call involving terminals with more sophisticated capabilities.

3. MCU AND TERMINAL CAPABILITIES

The capabilities of an audiovisual terminal are described in Rec H.242, and those of an MCU in Rec AV.231. It is clear that in a conference call the communication possibilities may be limited by the MCU as well as by the mutual capabilities of the participating terminals. The choice of MCU, at the time of access or prior reservation, must be such as to minimise such limitation; particular examples are:

(a) for a videophone conference call, the MCU must have a video processor (Rec AV.231), and an appropriate access (1B, 2B, H0 etc);

(b) for an audiographic conference call, the MCU must have an MLP processor.

4. PROCEDURES FOR CALL CONTROL

4.1 "Meet-me" Type - All Participants Dial In

(a) Convenor reserves MCU of appropriate capacity and capability (see section 3).

(b) Convenor notifies participants of date, time, duration, and number to dial.

(c) All participants dial in at or just after the reserved start time.

(d) The MCU auto-answers each in turn with message M1.

Case 1: Simple Terminals and Basic MCU

(e1) MCU connects each terminal into the conference and uses M2(n) via the audio mixer to announce the number of terminals now participating: this enables the convenor to check that no unauthorised terminals have dialled in.

(f1) Whenever there is a change to the number of active ports, M2(n) is repeated with the new value of n.

Case 2: MCU and Convenor Terminal Have MLP Capability

(e2) M1 is followed by M3, if the convenor has not yet dialled in. (NB: if the MCU has not been preset to expect a convenor with MLP, M3 is omitted and the conferees are added into the conference.)

(f2) When the convenor dials in, his terminal sets up MLP channel to the MCU (see Rec H.242, Section 9.2), which returns information as to the terminal numbers assigned to conferees who have already dialled in and are waiting.

(g2) The convenor selects one terminal, is connected by audio to it, greets the conferee, and transfers the connection to the conference proper (ie, via audio mixer and video/data if relevant); he repeats this until all are connected. He may also reject a caller if so desired.

(h2) The MCU keeps the convenor's terminal informed, via the MLP, of the status of its ports. At any time, the convenor may disconnect a terminal from the conference, and either speak directly to that conferee or clear down his call altogether.

Case 3: MCU and Some or All Terminals Have MLP Capability

(e3) Each MLP-capable terminal establishes a dialogue with the MCU, assigning terminal numbers, location and/or conferees' names (and perhaps associated microphone identifiers) and so on.

(e4) The terminals may be connected immediately into the conference proper, without waiting for the intervention of the convenor, or the MCU may be set to preclude this if specified at the time of booking.

(f3) The convenor is identified by insertion of a password obtained at the time of booking. The convenor may disconnect terminals from the conference, as in (h2) above; he may take later calls to the conference, with or without adding them in. The convenor may release his role to another conferee by passing him the convenor "token", using the MLP.

Messages

M1: This is an automatic conference control unit.

M2(n): The number of terminals now connected (to the conference) is n.

M3: Please wait to speak to the convenor.

M4: Your terminal is number 1; other terminals will be numbered in the order they are connected (to the conference).

M5: To add another conferee at any time, please key # followed by the (long-distance) code.

M6: Please now begin calling the other conferees.

M7(n): Terminal n has just been disconnected.

Note: experienced users may be annoyed if they have to wait for the messages M4-M6 to be completed before they can continue; therefore the system should be such that no waiting is necessary, and messages are cut off as soon as the convenor begins.

4.2 "Dial-out" Type - The Convenor Dials All Conferees

(a) The convenor identifies (via the service provider) an MCU of appropriate capacity and capability (see section 3) and reserves it if necessary.

(b) The convenor notifies participants of date, time and duration.

(c) The convenor dials the MCU at the appointed time.

Case 1: Simple Terminals and Basic MCU

(d1) The MCU auto-answers with message M1.

(e1) Messages M4, M5, M6 follow.

(f1) The convenor dials the conferees in turn.

(g1) On answering, each conferee is immediately connected into the (audio bridge or) conference proper; he identifies himself as he would for a telephone call.

(h1) If the convenor is not satisfied as to the suitability of a conferee to participate, he must key something unmistakable (such as three or more*) followed by the terminal number.

(i1) If for any reason a terminal is dropped from the MCU, the message M7(n) is heard.

Case 2: MCU and Convenor Terminal Have MLP Capability

(d2) The MCU auto-answers and an MLP channel is set up (Rec H.242, Section 9.2); instructions are sent to the convenor to guide him through the process of dialling the other conferees, following the same routine as for Case 1 above.

(e2) The convenor may choose to speak to called conferees individually before adding them to the conference proper.

4.3 Supplementary Service to a Point-to-Point Call

(a) Terminal A, speaking to B and wishing to add C and D, first forces Mode 0 (see Rec H.242, Section 6.2) and suspends the call to B.

(b) A dials an MCU of dial-out type and suitable capacity (see Section 3) - this may involve keying "Conference" or another combination if the MCU is attached to the local exchange; if A and the MCU both have MLP capability, this channel may be established to facilitate further action, otherwise message M1 is heard.

(c) Terminal A or its local exchange must detect the D-channel message completing the MCU-A connection, and must then forward the terminal-B connection to a second port of the MCU capable of auto-answer (despite the MCU being "dial-out" type).

(d) Both terminals now hear messages M4-M6 but, to avoid the possibility of contention, only the first to key # is acted upon; the other is ignored.

(e) The conference call proceeds as in Section 2 (d1) to (i1) above.

4.4 Operator-Controlled MCU

To be completed.

5. NETWORK IMPLICATIONS

The design of the general audiovisual system is such that, despite the wide variety of applications foreseen, the special requirements of the network are kept to a minimum.

5.1 HLC and LLC

It is a prerequisite that, where a connection between two equipments is requested, it should be completed. In the general case, when a user (A) calls another (B) he may be unaware of the exact type of audiovisual terminal possessed by B; likewise B is unlikely to decline to answer simply because A's audiovisual terminal is different from his own. In this context "audiovisual" should be taken to include audio-only terminals - narrowband and wideband telephones.

In the same way, the MCU cannot be aware of the exact types of the audiovisual terminals it is calling, in the dial-out case, nor of those dialling in, in the meet-me case.

There is therefore a strong case for the use of a generalised audiovisual HLC in the call-control process, together with LLC of "unrestricted digital" (or perhaps "7kHz bearer" if this solves the problem of interworking with PSTN telephones).

5.2 Supplementary Service

A method is required whereby a caller connected to another party is able to:

(a) key "M" to suspend his connection to the other party and either (i) be automatically connected to a suitable MCU, or (ii) dial/autodial the number of a suitable MCU;

(b) when the MCU is connected, either (i) he keys "Z" or (ii) his terminal recognises that the MCU has answered and acts as if "Z" had been keyed;

(c) the local exchange recognises the "Z" command, and re-routes the original called party to another port of the MCU;

(d) the caller then keys "X" to suspend his connection to the MCU, holding it at the local exchange;

(e) he dials the number of another party;

(f) when the other party answers, either (i) they speak first, and the caller then again keys "Y", or (ii) the calling terminal recognises that the other party has answered and acts as if "Y" had been keyed, without waiting for either party to speak;

(g) the local exchange recognises the "Y" command, disconnects these two parties from each other and connects both to the MCU (the original caller must be restored to the same MCU port as before).

It may be noted that b(ii) has the advantage of greater simplicity for the user. The choice of keys ~~X and Y~~ M x Y Z and of b(i)/(ii) are a matter for the terminal designer.

6. REQUIREMENTS ON IN-BAND SIGNALLING

6.1 "Meet-me" Type MCU

Using MLP a wide variety of procedures is possible.

Basic terminals and MCU do not have this capability. The use of recorded/synthesised speech for status indications from the MCU obviates the need for defining more C&I symbols, and for terminals to recognise and use them.

6.2 "Dial-out" Type MCU

The dial-out MCU could be designed to set up calls to other terminals from each of its ports, but would need to obtain from the convenor the numbers to be called. This is no problem if the convenor and MCU have MLP capability, but poses a problem if the convenor has only a basic terminal (not specifically designed for conferencing). Part of the procedure of section 5.2 would be applicable (points (d) to (g) but may be difficult to implement when the MCU is not attached to the convenor's local exchange.

Other possibilities are:

(a) Speech recognition in the MCU, to accept a verbal input for numbers to be called;

(b) A basic terminal may be designed such that, when a call is in progress and H.221 framed transmission is being used, keying 0-9, #, * results in the transmission of a corresponding C&I symbol, to be specified in Rec H.230.